**AUGMENTED REALITY FOR TRAINING FORMATION FLIGHTS: AN ANALYSIS OF HUMAN FACTORS**

Diego Hernandez Arjoni1, Ivan de Souza Rehder1, José Márcio Pereira Figueira2, Emília Villani1

1 Instituto Tecnológico de Aeronáutica, Brazil

2 Instituto de Pesquisas e Ensaios em Voo, Brazil

**Abstract**

Pilot training has been, for decades, aided by flight simulators with different characteristics and degrees of fidelity. However, many studies indicate that, despite the recognized contribution of simulator training, actual flying practice is still necessary, depending on the trained task. Some researchers show that the pilot's psychophysiological measures undergo alterations from the real flight to the simulated flight, affecting the transfer of training and reinforcing the need to maintain, to some extent, the training in the real flight. Considering the above, this work addresses the problem of basic formation flight training. This flight requires the use of at least two aircraft, resulting in high operational costs and risk of collision between aircraft. One way to overcome this problem is using augmented reality in flight, with the objective of replacing the leader aircraft with a virtual representation of it and providing a scenery close to real flight, but with reduced cost and risk of collision. As a scientific contribution, we investigate the benefits and limitations of training using augmented reality in a simulated environment, when compared with the usual training in a static simulator for the flight scenario in basic formation flight.

Keywords: augmented reality, human factors, flight simulator, in-flight training.

# Introduction

The safe and efficient training of pilots has been a challenge since the dawn of aviation. It has motivated the research and development of flight simulators, with different features and degrees of fidelity. Despite the recognized contribution of simulator training, published works indicate that training in real flight is still necessary (ROESSINGH, 2005). Previous studies indicate that the pilot’s psychophysiological and endocrine reactions in a real flight are different from in a simulated flight (VELTMAN, 2002), (JOHANNES et al., 2017) and (LEINO, 1999). Also, results obtained in experimental works point to higher levels of stress and workload in a real flight. They also show that the level of stress influences the training efficiency (MCCLERNON, 2009), (MCCLERNON, MCCAULEY, et al., 2011). The identification of limits and benefits of different training environments remains an open problem, whose answer depends on a series of factors, such as environment conditions, pilot’s experience, manoeuvre, among others.

Traditionally, training is carried out in two phases: simulator training and in-flight training. The limits of the use of simulators for training has been a constant topic of research (VALVERDE, 1973), (HAYS, JACOB, et al, 1992), (ORITZ, 1993) (BLUME, 2010)). The several reviews carried out in the literature indicate that the transfer of training depends on the trained task, the training environment and the method used for evaluation (HAYS, JACOB, et al., 1992), (BELL and WAAG, 1998). Particular attention has been devoted to investigations into the need to integrate a motion system into the flight simulator and the influence of the sensation of motion on pilot learning and behaviour (BÜRKI-COHEN and GO, 2005), (BÜRKI-COHEN, SPARKO, 2005) (BLUME, 2010).

The advent of digitization brings in new technologies, such as virtual and augmented reality, that can be used to create new training environments (LABEDAN et al., 2021), (AUER et al., 2021), (LEE, 2016). Among the new options brought by digitalization, the one analysed in this work corresponds to the use of augmented reality in a real flight to create a training scenario. In this case, elements external to the aircraft are visualized by the pilot during the flight using an augmented reality device. This proposal provides flexibility and diversity for the creation of scenario configurations and can be used for different types of training, such as reconnaissance of aircraft in an intercept mission, search and rescue, air-to-air combat, air-to-ground combat, close air support, among others.

The use of augmented reality systems in aircraft is not new. Head up displays and helmets are commercial products extensively used in aviation (ARTHUR, 2015; ARTHUR, 2017; FOOTE et al 2015). However, in this paper we focus on the use of augmented reality for reproducing elements and objects that would be part of the outside environment, in order to compose a training scenario.

Figure 1 compares the training environment investigated in this paper (real flight with augmented reality) with the traditional training environments (simulator and real scenario), according to their risk, cost and fidelity.

*Increasing risk, cost and fidelity*

Training in a simulator

Training in a real flight with augmented reality scenario

Training in a real scenario

Figure 1. Different training environments.

The training scenario we investigate in this work is that of military formation flights. Formation flight training requires at least two aircraft, imposing additional costs and logistical limitations. Moreover, this type of flight presents additional risk due to the proximity between aircraft, as evidenced in mid-air collision accidents (MCSPADDEN, 2019), (BCAA, 2016), (DICKSTEIN, 2019).

Briefly, in a formation flight of two aircraft (leader and wingman), we investigated the possibility of training the wingman using augmented reality to show the leader aircraft. Our premise is that this kind of training can replace, to some extent, the training with two real aircrafts.

More recently, among the initiatives to use augmented reality systems in a real flight are the solutions developed by the Netherlands Aerospace Laboratory (NLR) with focus on air combat training. They propose the superposition of real and virtual environments to simulate Beyond Visual Range (BVR) combat situations (ROESSINGH, VAN SIJLL and JOHNSON, 2003), (KRIJN and WEDZINGA, 2004). An extension of the embedded training concept to visual combat situations is identified as a challenge yet to be solved (LEMMERS, 2008), (TASK GROUP HFM-165, 2014). This challenge is receiving attention and efforts from companies such as RED6, which proposes the superposition of virtual 3D elements to the pilot's field of view during a real flight (ROBINSON and BICANIC, 2019).

The works developed at NLR and RED6 indicate that any technological bottlenecks associated with the use of augmented reality for in-flight training should be resolved in the near future. However, in both cases, we observe a lack of a systematic evaluation of the impact on the pilots when real objects are replaced by virtual ones, regarding factors such as workload, stress and performance. In addition, the case of formation flight is not addressed in any of these works.

Assuming that the transfer of training from the simulator to a real flight is not complete and imposes the need to maintain, to some extent, the training in real flights, we propose following hypothesis: *“in the initial stages of formation flight training, the projection of the leader using augmented reality technology provides a similar learning to that obtained with the visualization of the real aircraft.”*

This hypothesis is investigated through experiments carried out using a proof of concept of an augmented reality system integrated in a high-fidelity simulation environment. Although the validity of the results obtained in the simulation environment is limited, it is considered an essential step in the feasibility analysis of training based on the use of augmented reality in flight.

Despite the apparent contradiction, once that it is exactly the difference between the real flight and the simulated flight that justifies the proposal to use augmented reality in flight for training purposes, we considered that several aspects of the replacement of the real leader aircraft by a projection in augmented reality can be evaluated in a high-fidelity simulation environment, albeit in a limited way and requiring subsequent in-flight confirmation.

Summarizing, we perform a set of experiments with the use of augmented reality for formation flight training, where the leader aircraft is replaced by a virtual projection. The experiments are designed to investigate the following human factors: workload, stress, performance and training.

The paper is organized as following. Section 2 presents a review of the literature and discusses researches that investigate the transfer of training between simulated and real flight. It addresses works that compare the pilot’s psychophysiological measurements obtained in real flight with those obtained in simulated flight. It also describes initiatives related to training in real flight with virtual elements. In Section 3, we present the prototype developed at the Aeronautics Institute of Technology for integrating an augmented reality system with an existing high-fidelity simulator. Then, Section 4 proposes an experiment to evaluate the impact on the pilot of the use of augmented reality. The results obtained in the experimental campaign are discussed in Section 5. Finally, Section 6 summarizes the main conclusions and discusses future work.

# Literature Review

The effectiveness of training in different environments have been extensively discussed in the literature, with different results and conclusions. Reid (REID, 1975) presents one of the first studies comparing the transfer of training obtained from simulated flights with that obtained from real flights. He discusses the efficiency of training on a fixed-base flight simulator composed of a simplified cockpit of a T-38 aircraft with the image of the leader aircraft is projected on the spherical screen (WOOD, HAGIN, et al., 1972). Reid’s experiment was performed with 70 student pilots distributed in 3 groups: the first group trained on the simulator, the second group trained on the real T-38, and the last group was not subjected to any training. Following the training rounds, the pilots were evaluated by flight instructors. The results shows a significant difference between the trained and untrained pilots, but not between those trained in real and simulator flights, suggesting both are equivalent. Roessingh (2005) evaluates the transfer of piloting skills acquired in a low-fidelity environment to real flight, for the case of acrobatic flights. The experiment was carried out with 21 student pilots with no experience in this type of manoeuvre. The pilots were separated into 3 groups: the first one was trained in real flights, the second one received the same training of the first group plus additional training in a high-fidelity simulator, and, for the third group, the high-fidelity simulator was replaced with a low-fidelity one. The pilots’ flights were evaluated in two ways: objective performance metrics based on aircraft data and scores given by instructors. In the opposite direction of the Reid’s work, the results do not indicate a significant difference among the groups, i.e., no contribution resulting from simulator training is identified, highlighting the importance of real flight training.

Among the works devoted to analyse similarities and differences between simulated and real flight, Leino (1999) compares the pilots’ endocrine reactions resulting from flights in the BA Hawk Mk 51 aircraft and its training simulator. The experiment was performed with 5 experienced pilots and 5 novice pilots from the Finnish Air Force. Blood samples were collected 30 minutes before and 10 minutes after the flights and tested for the various hormones. The results show significant variation in the prolactin and atrial natriuretic peptide (ANP) hormones from real to simulated flights, suggesting an increase in stress associated with real flight. Veltman (2002) analyses the sensitivity of various psychophysiological indicators to real and simulated flights, in an experiment with 20 pilots from the Dutch Air Force. The pilots performed 6 tasks in the Frasca 141 flight simulator and in real flights with the Slingsby aircraft. The results indicate an increased level of the cortisol hormone after the real flight, which did not happen after the simulated flight, confirming the conclusion of Leino that the real flight is more stressful. Johannes (2017) proposes a metric, called PAV (Psychophysiological Arousal Value), that combines different psychophysiological measures into a single index. PAV was used in an experiment where 15 pilots (5 novices and 10 experienced) had to perform tasks with different levels of difficult during real and simulated flights. Results show statistically significant difference in the PAV metric for real and simulated flights for the case of novice pilots performing the most difficult tasks. However, no difference is identified for experienced pilots.

Following a different approach, McClernon (2009) analyses the influence of stressors on pilot training. He performed a comparative experiment using 29 novice pilots, divided in 2 groups. Both groups were submitted to training in a flight simulator. During the training, one of the groups was exposed to a stressor that consisted of submerging one of the participant's feet in water at 9°C for a period of 10 to 15 minutes. Finally, evaluation flights were carried out, with both groups, under the effect of the stressor. The authors collected data from ECG, GSR, performance and subjective stress ratings. The results show a significant influence of training conditions on pilot transfer of training. The group exposed to stressors during training performed better than the group that had to deal with the stressors for the first time during the assessment flight. In 2010, the authors performed a new experiment, where the they maintained the same configuration of the previous experiment, but in this case the transfer of training was assessed in real flight. Again, the results indicated a greater transfer of training for the group subjected to training under the effect of stressors, confirming the importance of real flight training.

Summarizing the works discussed above, differences between the pilots’ psychophysiological and endocrinological reactions in real and simulated flight indicates greater stress and workload in the real flight. Moreover, studies also indicate that training under stress may improve performance in the real flight. These two conclusions are consistent with the works that identify an improvement in the pilots’ performance when training in real flights. These works motivate the proposal of this paper, which proposes the use of augmented reality for creating training scenarios in a real flight.

The idea of developing solutions to create training scenarios in real flights is not new. In 2003, Roessingh et al. performed a study to identify the necessary requirements for an embedded combat training system. The architecture proposed by Roessingh et al was later used in the development of the E-CATS system, an embedded training system capable of simulating combat situations outside the visual field, air/air and air/ground (WEDZINGA, 2005), (KRIJN and WEDZINGA, 2004). Later, an evolution of the E-CATS system that supports situations with two or more airplanes in out of visual range combat was presented in (LEMMERS, 2008) and (ROESSINGH and VERHAAF, 2009).

More recently, a number of works discusses the challenges identified within the framework of a NATO (North Atlantic Treaty Organization) research group regarding the development of embedded training systems and embedded virtual systems (MAGEE, SOTTILARE and ROESSINGH, 2011), (ALEXANDER, SOTTILARE, et al., 2012) and (TASK GROUP HFM-165, 2014). Among other recommendations, these works highlight the need of developing solutions for within visual range combat. These technological challenges have been explored by companies such as RED6, which proposes an augmented reality system called A-TARS (Airborne Tactical Augmented Reality System) integrated into a helmet mounted display. Aligned with the proposal of this work, A-TARS generates holograms of external elements in the execution of a combat flight. However, different from our work, the focus is on combat situations, while our application is devoted to training formation flights. Moreover, none of their published works investigates, in a systematic way, the impact of replacing real elements with virtual ones.

# Evaluation Prototype

In order to investigate the impact of replacing the real leader aircraft with a virtual one when training formation flights, we developed a prototype of an augmented reality system to be tested in a high-fidelity flight simulator. The prototype should allow experiments to be carried out in an initial stage of feasibility studies, based on the premise, widely adopted in Aeronautics, that experiments carried out in a simulation environment allow us to increase our knowledge about a new technology and how this technology interacts with the human being, thus contributing to elevate its Technology Readiness Level (TRL), despite the differences between the simulation environment and the real flight.

The purpose of the evaluation prototype is to anticipate the identification of drawbacks and limitations that arises when replacing training scenario (a) with (b), where:

1. Formation flight where the leader aircraft is a real aircraft;
2. Formation flight where the leader aircraft is a projection of augmented reality glasses integrated into the wingman aircraft.

Once that the comparison of training scenarios (a) and (b) is not feasible at early stages of development, we propose the following comparison, to be performed in SIVOR, a high-fidelity simulator:

1. Formation flight where the leader aircraft is projected by the visual system of the SIVOR simulator;
2. Formation flight where the leader aircraft is a projection in augmented reality glasses integrated into the SIVOR simulator.

The transposition from “comparison (a) versus (b)” to “comparison (c) versus (d)” considers that, from the perspective of the pilot, the leader aircraft, when projected by the visual system of the SIVOR simulator (situation (c)) is perfectly integrated to the external environment and, therefore, corresponds to the situation where the aircraft is real (training scenario (a)). Likewise, situation (b), where the leader is a holographic projection in the real world, can be understood as equivalent to situation (d) where, where the leader aircraft is also projected through AR glasses.

SIVOR is a robotic flight simulator developed at ITA. It was conceived as a level-D full flight simulator [FAA, 2016]. The cockpit of SIVOR, based on that of a real aircraft, is illustrated in Figure 1. The SIVOR motion system, which uses an anthropomorphic industrial manipulator, is not used in this work. Figure 3 illustrates the execution of a formation flight in the training scenario (c), where the leader aircraft is projected by the visual system of SIVOR.

Televisão ligada ao fundo

Descrição gerada automaticamente com confiança média

Figure 2. Cockpit of SIVOR flight simulator.

Monitor de computador

Descrição gerada automaticamente com confiança baixa

Figure 3. Running training scenario (c), with the leader aircraft projected by the visual system of SIVOR.

Finally, Figure 4 presents the leader aircraft projected by the augmented reality system next to the leader aircraft projected by the visual system.

Imagem de vídeo game

Descrição gerada automaticamente com confiança média

Figure 4. Comparison of the leader projection in scenarios (c) and (d).

The main components of the SIVOR simulator and the evaluation prototype are illustrated in Figure 5.

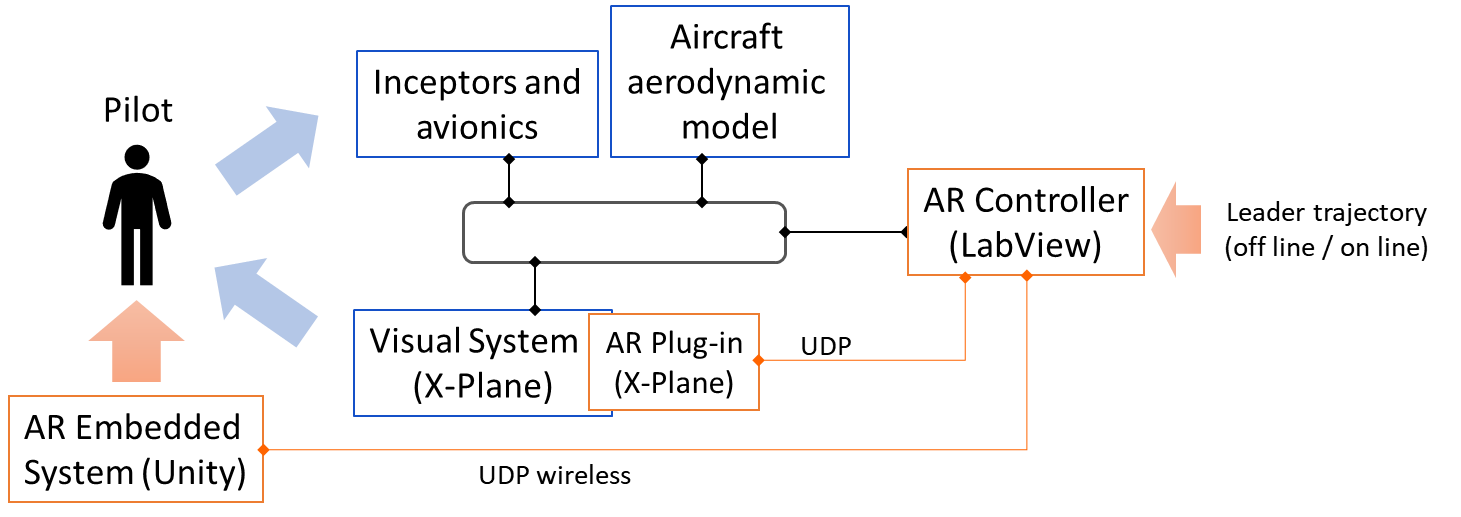


Figure 5. Evaluation prototype integrated into the SIVOR infrastructure.

In the SIVOR simulator, the pilot provides commands through the inceptor and avionics devices. The pilot inputs are transmitted to the aircraft aerodynamic model that calculates the expected behaviour of the aircraft for those inputs. The outputs of the aerodynamic model, such as the aircraft attitude and position, are provided to the visual system. The SIVOR visual system is based on the commercial tool X-Plane. It defines the image of the outdoor environment and projects it on a 180o screen that the pilot sees through the cockpit window. As part of SIVOR infrastructure, a real-time network integrates the inceptors, avionics devices, aerodynamic model and visual system, assuring the consistence of the simulation.

The evaluation prototype is composed of three additional modules that are integrated to the SIVOR simulator: AR embedded system, AR Controller, and AR Plug-in. The AR Embedded System was developed in Unity and is loaded onto the augmented reality device. It projects the leader aircraft in the case of training scenario (d). The AR plug-in is added to the visual system, with the function of configuring and controlling the visualization of the leader aircraft in the case of training scenario (c). Finally, the AR Controller is developed in LabView and is responsible for receiving the data corresponding to the trajectory of the leader aircraft and providing it to both the AR Plug-in and AR Embedded System. In order to provide flexibility for future works, the evaluation prototype allows the leader trajectory to be defined in different ways, such as: reproduction from pre-recorded flight (playback file), real external pilot in-the-loop, autopilot defined by artificial intelligence routines, among others. Following, we describe in detail the three components of the evaluation prototype.

## AR Embedded System

The AR Embedded System corresponds to the formation flight application embedded in the HoloLens device and used in training scenario (d). It was developed in the Unity 3D platform and has two main functions: positioning elements in the Unity world and controlling the leader aircraft projection. Following we describe the options considered in the implementation of each function, as well as the adopted solution.

*Positioning of elements in the Unity world*

For a formation flight training session, the AR Embedded System must position two types of elements:

* Active or dynamic elements, which correspond to objects that moves in the Unity world, such as camera, leader airplane, virtual lights; and
* Passive or static elements, which correspond to objects that do not move in the Unity world, such as the cockpit occlusion model.

The inertial reference system of the Unity world is attached to the cockpit of the wingman aircraft. Due to the relative motion between the wingman aircraft and the ground, the ground is in motion with respect to the inertial frame and is, therefore, an active element. If, on the contrary, the inertial frame were positioned on the ground, the wing aircraft would be in motion and would be an active element.

Furthermore, the augmented reality application must consider interaction between real and virtual objects. The formation flight application requires the implementation of occlusion routines to preserve the simulation fidelity. These routines are responsible, for example, for not allowing the visualization of the parts of the leader aircraft that would appear over the cockpit panel, as illustrated in Figure 6, providing the user with an idea of depth and increasing the realism of the simulation.

Texto

Descrição gerada automaticamente

Figure 6. Visualization of the leader aircraft without (on the right) and with (on the left) the aircraft occlusion by the cockpit.

In order to project the active and passive elements, it is necessary to define the position of the pilot, who is wearing the augmented reality glasses, in the Unity world. The element corresponding to the pilot in the Unity world is the camera. The three most common methods to resolve this issue are:

* Standard positioning method: the Unity world, by definition, uses the position of the glasses at the start of the application as the origin or 0 point of the inertial reference frame of the Unity world, fixing the direction of the gravity vector as the z-axis. From that moment on, the update of the position of the glasses, that is, of the pilot's head, is performed based on the data from the inertial sensors embedded in the glasses.
* Environmental recognition method: HoloLens sensors perform a mapping of the environment that results in the identification of entities such as horizontal planes, vertical planes, cylinders, cones, among others. These entities are used to create a 3D model of the environment. The holograms are then positioned in relation to some of identified elements.
* Marker recognition method: consists of identifying pre-programmed images using the glasses camera and pattern recognition routines. After identifying the marker, the relative position between it and the glasses is calculated to define the position of the inertial reference frame of the Unity world.
* Manual calibration method: Hololens allows user-supplied inputs to adjust the positioning of the glasses in the Unity world.

During the development of the AR Embedded System, the marker recognition method was initially considered, due to its simplicity. We used the Vuforia software to program a QR code as the target image. When the QR code is recognized, its position is assigned to the camera of the Unity world, that is, to the glasses. The corresponding overlay of the aircraft image to the real world, i.e., the view of the leader aircraft from the SIVOR cockpit, is shown in Figure 7 for level flight and a curve flight.

Uma imagem contendo Diagrama

Descrição gerada automaticamente Diagrama

Descrição gerada automaticamente

Figure 7. Visualization of leader aircraft during a level flight (on the right) and a curve flight (on the left) using augmented reality glasses.

After a testing campaign, the marker recognition method was abandoned for the following reasons:

* The system has high latency and abrupt changes to the leader aircraft's positioning often occur during flight.
* The dark environment of the cockpit and the low resolution of the glasses camera make the QR code recognition very unstable.
* The high energy consumption of the HoloLens camera, which must remain on constantly, shortens battery life, which could make longer training impossible.

The positioning approach based on environment recognition was also discarded after some tests due to the difficulty of detecting elements in an environment as restricted as the interior of the cockpit. In addition, the use of this functionality for occlusion also ends up being unfeasible due to large variations in the 3D mesh, which ends up interfering significantly in the fluidity of the projection of the leader aircraft.

As a final solution for the positioning of the elements, a hybrid approach is adopted that combines standard positioning with manual calibration methods. This approach starts from a standard initial positioning, approximately corresponding to the position of the pilot's head in the cockpit. Then, a calibration is performed based on directional and guidance commands that the pilot provides to the HoloLens device using a bluetooth keyboard.

For the calibration step and the leader occlusion, a 3D model of the cockpit was created in the Unity world and is used as a mask. This model was obtained from the scanning of the internal part of the flight simulator, merged with a 3D design model. The steps of creating the 3D model of the cockpit are illustrated in Figure 8. Figure 9 shows the cockpit from the pilot's perspective.

Brinquedo de lego

Descrição gerada automaticamente com confiança média

Figure 8. From left to right: scanned model, model composed of separate console and cockpit parts; final 3D model.

In the AR Embedded System, all objects in the Unity world are hierarchically organized as 'child objects' of a single 'parent object' called the 'AR Universe'. In such a hierarchical relationship, when the 'parent object' has its position or orientation modified, the same occurs with the 'child objects', in such a way as to keep unchanged the relative positioning between 'child objects' and 'parent object'. When the 'child object' is moved, the position and orientation of the 'parent object' remains unchanged.

Once the pilot is in the cockpit of the SIVOR simulator, the AR Embedded System is initialized and executes the calibration routine. The inertial frame of reference is assigned to the current position of the HoloLens and all elements of the Unity world are positioned at an approximate position. At this point, the pilot sees the 3D model of the cockpit in red, close to its final position. Using a set of commands, the pilot is able to make small adjustments to the position of the “AR Universe” and, consequently, move all the elements of the virtual world. The pilot is instructed to follow this process until the red virtual cockpit perfectly overlaps the real SIVOR cockpit, correcting the relative position between the real world and the virtual world. Figure 6 22 illustrates the virtual cockpit model (in red) overlaying the real cockpit during the manual calibration process. Finally, the color of the cockpit is changed to black and it works as an occlusion mask. Once the calibration is complete, the AR Embedded System is ready to begin the formation flight training.

Ícone

Descrição gerada automaticamente com confiança baixa Uma imagem contendo água, avião, mesa, montanha

Descrição gerada automaticamente

Figure 9. Overlapping between the calibration mask and the real cockpit: central column (on the left) and aircraft window (on the right).

*Controlling the leader aircraft*

A first approach implemented in the evaluation prototype for controlling the leader aircraft consisted of calculating the relative position of the leader aircraft in relation to the wingman aircraft using coordinate transformations to define the position and orientation of the leader aircraft in the coordinate system attached to the wingman. The necessary calculations were implemented in the AR Controller. However, this approach resulted in a significant latency, compromising the quality and fidelity of the simulation.

As an alternative solution, the project team adopted an approach that was implemented in the AR Embedded System and uses the camera of the Unity world and hierarchical relationships between objects. In an augmented reality application, the camera of the Unity world is linked to the glasses’ sensors that capture the user movement. The camera position and orientation in the Unity world can be automatically updated as the user moves, ensuring that the projected image is updated. In the proposed approach, the position and orientation of the camera in relation to the AR Universe is not modified or affected by the relative distance between the leader and wingman, but is a function only of the position of the pilot’s head in the cockpit. As a consequence, it is controlled directly by HoloLens. Moreover, the AR World is a child element of the AR Universe, and corresponds to the ground. Instead of the AR World remaining fixed and the wingman moving relative to it, the aircraft remains fixed and the AR World moves in the opposite direction of the aircraft's flight. The leader aircraft is also defined as a child element of the AR World. When the AR World moves relative to the wingman, the leader aircraft also moves, automatically updating its position relative to the wingman aircraft.

Following this approach, the virtual world in Unity is composed of the following elements:

* AR Universe: empty object responsible for being a wrapper for all elements of the virtual environment except for the camera;
* AR World: empty object responsible for being the envelope of the entire mobile world;
* Rotational World: empty object responsible for executing the inverse rotations of the wingman aircraft;
* Translational World: empty object responsible for executing the inverse translations of the wingman aircraft;
* Leader Aircraft: 3D object corresponding to the leader aircraft;
* Cockpit: 3D cockpit-shaped object that works as an occlusion mask for the AR Embedded System;
* Camera: object that represents the pilot's head position in the Unity world.

We observe that retrograde movement strategies, where the world moves instead of the user, are commonly adopted in games, as it considerably reduces the computational load associated with scenario generation and rendering.

## AR Controller

The AR Controller is responsible for integrating and synchronizing the data exchange between the AR Embedded System and the other components of the SIVOR flight simulator. It was developed in LabView, with the following functions:

* Communicate with the real time network of SIVOR flight simulator to receive the wingman flight data;
* Communicate with the AR Plug-in and AR Embedded System to send the wingman and leader flight data;
* Provide different mechanisms for generating the leader flight data, such as playback a recorded flight, run routines that generate the flight from a set of waypoints, or integrate the evaluation prototype with another simulator;
* Provide the human-machine interface (HMI) for configuring and managing the formation flight simulation.

## Plug-in

The AR Plug-in was developed using the X-Plane SDK (Software Development Kit) provided by Laminar Research. Its execution is integrated and coordinated by the X-Plane software tool.

The AR Plug-in is used in the case of training scenario (c). Its main functions are: enable communication and data exchange between X-Plane and the AR Controller using UDP protocol, create and configure the leader aircraft, execute the coordinate transformation routines, and update the position of the leader aircraft image.

The prototype describe in this section was used in an experimental evaluation campaign described in the next section. Results are discussed in Section 5.

# Evaluation experiment

The evaluation experiment aims at comparing scenarios (c) and (d). We want to verify whether or not the projection of the leader using augmented reality technology provides the pilot with a similar experience to that obtained with the visualization of the leader aircraft in SIVOR visual system.

The evaluation focuses on three issues:

* the pilot performance, i.e., whether or not a training environment is easier than the other;
* the training effectiveness, i.e., how the pilot performance evolves with training;
* the pilot workload and stress in each scenario.

The performance is considered as the ability to maintain a relative distance between the leader and the wingman. Over a period of time, performance is determined by the mean and variance of the relative distance between the airplanes, and also the amount of time that the pilot was able to maintain the relative distance within a specified acceptable tolerance. The training effectiveness is related to the improvement in the pilot performance after training.

The evaluation of workload and stress is performed using data from physiological sensors and a subjective workload scale (GAWRON, 2008). The physiological sensors are the electrocardiogram (ECG), which monitors the pilot's cardiac response, and the electrodermal activity (EDA) sensor, which shows how the skin conductivity is affected by the sweating activity (BRAITHWAITE, 2015, BOUCSEIN, 2012). The subjective scale is the NASA Task Load Index (NASA-TLX) questionnaire, one of the most commonly used subjective assessment methods (HART and STAVELAND, 1988; HART, 2006). The NASA-TLX is a multidimensional rating scale that combines six factors to estimate perceived workload: mental demand, physical demand, time demand, performance, effort and frustration. Each factor is evaluated on a scale with 21 levels. A binary classification between pairs of factors determines weights for each one of them.

## Hypotheses of the experiment

The expected result for the experiment is that the training effectiveness, the pilot performance and the pilot workload are not affected by the introduction of the augmented reality system. Therefore, the following hypotheses are evaluated in this experiment campaign:

* H1 – The pilot performance does not depend on the training scenario;
* H2 – The training effectiveness does not depend on the training scenario;
* H3 – The pilot workload and stress do not depend on the training scenario.

## Experiment procedure

This experiment was approved by an ethical committee according to the Brazilian regulation, as part of process CAAE 90300818.2.0000.5503.

The experiment was designed to be performed by amateur pilots, with different levels of experience of piloting in flight simulators with varying levels of fidelity. It was organized in four steps, as follows.

* *Step 1. Familiarization with the simulator*

The pilot undergoes the briefing procedure, signs a consent form, and receives an introduction about the SIVOR simulator and commands necessary for the experiment. Then, he/she performs a familiarization flight of 10 min, during which the pilot is instructed to perform some altitude and speed modification tasks.

* *Step 2.* *Training section*

This step follows a between-subjects configuration, where the pilots participating in the experiment are divided into two groups. Each group performs a training task in one of the evaluated scenarios (scenario (c): the leader aircraft is projected by SIVOR visual system; or scenario (d): the leader aircraft is projected by the augmented reality glasses).

The training task consists of:

1. Short flight during which the wingman autopilot is on. This flight places the leader and wing aircraft aligned and levelled, at a distance of 250m between the planes. Then, the autopilot is turned off and the pilot takes over the aircraft control. This step was introduced to reduce the difficulty of the experiment, avoiding the need for the pilot to perform the pairing between the aircraft.
2. Supervised training, carried out for a period of 15 minutes, during which the pilot receives instructions on how to improve his performance.
3. Unsupervised training, carried out for a period of 30 min.

The training consists of a basic formation flight, where the wingman must follow a leader aircraft that remains in constant heading, performing only small oscillations around an altitude of 10000 ft. Although it appears to be a simple task, maintaining the formation poses a high degree of difficulty for amateur pilots. In order to minimize the difficulty of the training task, the wing aircraft rolling freedom is limited, allowing only elevation control, through the sidestick, and power control, through the throttle. Furthermore, to provide a reference for the pilot, a green-yellow-red light indicates when the wingman-leader distance is below 250 m, between 250 and 500 m or larger than 500 m, respectively. The leader trajectory is a pre-recorded flight, guaranteeing that all the pilots are submitted to the same training task.

* *Step 3. Performance assessment*

In this step, unlike the previous one, the participants must perform formation flights using the two configurations (scenario (c) and scenario (d)). He/she performs 6 flights, 3 with each configuration. The order in which the scenarios are used is random. Each flight lasts 3 minutes and the task to be performed is the same as the unsupervised training in the previous step. Between two flights, the pilot has a 2 min break to relax.

* *Step 4. Conclusion*

After completing the formation flights of Step 3, the pilot is invited to answer 2 NASA-TLX questionnaires, one for each configuration.

## Data processing

In order to test the hypotheses presented in Section 4.1, the collected data are used to estimate performance, training effectiveness and workload.

***Performance Variables – Hypothesis H1***

The pilot performance at a given time interval is calculated from the GPS data (latitude and longitude) of the leader and wingman aircraft and consists of the following variables:

* Average of the distance between wingman and leader, in meters (*MeanDist*);
* Standard deviation of the distance between wing and leader, in meters (*StdDist*);
* Training time of during which the distance is below 250 m, in seconds (*T250*).

These variables are calculated for the 6 flights of Step 3, each corresponding to a 3-min time interval.

***Training Variables – Hypothesis H2***

In order to analyse the training effectiveness, for each pilot, the performance variables (*MeanDist*, *StdDist* and *T250*) are calculated at the following time intervals:

* INI: 1-min interval, starting 8 minutes after the beginning of the unsupervised training (Step 2);
* END: 1-min interval, before the ending of the unsupervised training (Step 2).

The difference in performance between the beginning and the end of the unsupervised training session is determined as the difference between the value of each variable in the two intervals, according to Eq. 1 to 3:

*DiffMeanDist = MeanDist(INI) – MeanDist(END)*(1)

*DiffStdDist = StdDist(INI) – StdDist(END)*(2)

*DiffT250 = T250(INI) – T250(END)*(3)

***Workload and Stress Variables – Hypothesis H3***

Two set of variables are used to estimate workload and stress: variables derived from NASA-TLX questionnaire and variables derived from the physiological sensors (ECG and EDA).

At the end of the experiment, each pilot answered two NASA-TLX questionnaires, one for each scenario. For each questionnaire the pilot attributes a score to each of the six factors that composed the NASA-TLX dimensions. Then, the pilot is presented with pairs of factors (e.g.: mental demand x effort) and has to select the one it considers more relevant. After the 15 comparisons, the weight of each factor is determined based on the number of times it prevailed and a weighted mean of factors is calculated.

In the experiment, each factor of NASA-TLX is considered as a variable, as well as the weighted mean of the six dimensions: mental demand (*MD*); physical demand (*PD*); temporal demand (*TD*); performance (*PF*); effort (*EF*); frustration (*FT*); and weighted mean of NASA-TLX dimensions (*NTLX*).

From the ECG signal, a number of variables can be calculated, both in time and frequency domain (MANISKKA, 2016, MARINESCU, 2018) (MACFARLANE, 2011) (HIDALGO-MUNOZ, 2018) (SHAFFER et al, 2017). This work considers the number of beats per minute (*BPM*) and the mean value of RR peak intervals (*RRi*) as time domain variables. These two variables are related to the heart rate. While the *BPM* is expected to increase with workload and stress, the *RRi* is expected to decrease. From the frequency domain, the ratio between the power of low frequency (0.04 - 0.15 Hz) and high frequency 0.15 - 0.4 Hz) components (*LF/HF*) of the heart rate variability spectrum is used. As the heart rate variability is expected to decrease with increasing workload and stress, the variable *LF/HF* should increase. Finally, the output of the EDA sensor is the mean value of the skin conductance response (*EDA* variable) and is expected to increase in a stressful situation. Summarizing, 4 variables are calculated from the physiological sensor measures: *BPM*, *RRi*, *LF/HF* and *EDA*.

The ECG and EDA data are collected using the CAPTIV software and exported. ECG data is analysed in Kubios platform, while EDA data are processed using an in-house Python routine.

# Results

The experiment describe in the previous section was performed with a set of 10 individuals with some experience of piloting airplanes in flight simulators. None of them was accredited pilot or had experience piloting real airplanes.

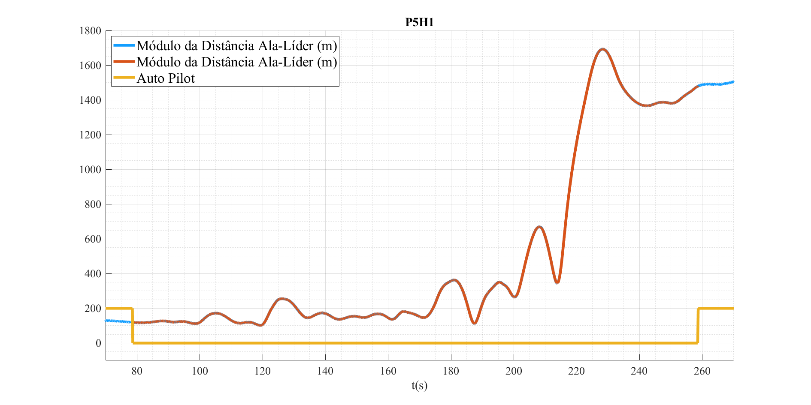
The individuals were divided in two groups. Following the procedure described in the previous section, each group trained flight formation in a different scenario (SV: the leader aircraft is projected by the visual system of the SIVOR simulator; AR: the leader aircraft is projected by the augmented reality glasses) and them performed the evaluation using both scenarios.

Due to the low number of samples and their expected heterogeneity, the statistical analyses described in this section, which are based on analysis of variance (ANOVA), consider a 90% confidence interval, equivalent to a p-value < 0.1.

## Performance analysis – Hypothesis H1

The performance evaluation is done with the data collected during the Step 3 of the experiment, when the pilot skills are tested in both training scenarios.

In order to illustrate the pilot output, Figure 10 brings, as example, the distance between the leader and the wingman for two 3-min flights. The flight on the left of Figure 10, is an example of low performance. It is possible to see that by the end of the interval the distance to the leader aircraft increases well above the target distance. On the other hand, the graph on the left of Figure 10 is an example a pilot with a good performance.

 Gráfico, Gráfico de linhas

Descrição gerada automaticamente

Figure 10. Examples of 3-min flight – distance between leader and wingman.

The first step of the data analysis was to calculate the performance variables for each 3-min flight: *MeanDist*, *StdDist* and *T250*. Considering that 10 pilots participate in the experiment and each of them performed 3 flights using each of the two training scenarios, a total of 60 values were obtained for each variable.

In order to evaluate the influence of training scenarios on these performance variables, ANOVA tests are performed using the statistical model of Eq. 4:

(4)

where:

* : output variable (*MeanDist*, *StdDist* or *T250*);
* : average of the output variable;
* : variance due to S factor, related to the training scenario;
* : variance due to P factor, related to the pilot;
* : variance due to the interaction between the factors S and P;
* : random error.

The P factor exposes eventual differences among the pilots, which could be due to a number of reasons, such as skill, experience, emotional condition, physical condition, among others. The S factor is related to the training scenario (VS: the leader aircraft is projected by the visual system of the SIVOR simulator; AR: the leader aircraft is projected by the augmented reality glasses).

Before applying ANOVA tests, the performance data were submitted to the Shapiro-Wilk normality test and the Bartlett homoscedasticity test. However, in the case of the variables *MeanDist* and *StdDist*, the p-value obtained from the Shapiro-Wilk test did not confirm that the data followed a normal distribution. A qualitative analysis of the flight showed that when the pilot error increases, the pilot efforts to compensate it may cause increasing oscillations in the distance between the airplanes, similar to what occurs in Pilot-Induced Oscillations (PIO) well-known phenomenon. In order to compensate for this exponential effect, the variables *MeanDist* and *StdDist* were submitted to a logarithm transformation (*LnMeanDist* = ln(*MeanDist*) and *LnStdDist* = ln(*StdDist*)) and the resulting variables were used instead. This transformation was sufficient to bring the data within the level of confidence adopted in this work (p-value > 0.1), as presented in Table 1.

Table 1. Analysis of normality and homogeneity of variance – Performance variables.

|  |  |  |  |
| --- | --- | --- | --- |
| **Test** | **LnMeanDist**  **p-value** | **LnStdDist**  **p-value** | **T250**  **p-value** |
| Shapiro-Wilk | 0.2086 | 0.4122 | 0.3103 |
| Bartlett P | 0.4155 | 0.4899 | 0.1087 |
| Bartlett S | 0.9111 | 0.3645 | 0.5045 |

Following, the data were submitted to ANOVA in order to check the influence of each factor, as well as the interaction among them. The results are presented in Table 2. According to the criterion previous presented, p-value < 0.1 indicates that the corresponding factor has a significant contribution to the output variable. As expected, the pilot (P factor) is significant for all the three variables. Its influence is due to the differences in previous flying experiences and natural abilities of the pilots. The training scenario (S factor) had a significant effect only for the variable *LnMeanDist*. The introduction of the augmented reality system resulted in the pilot increasing the distance from the leader aircraft. Finally, the interaction among factors was significant for both variables *LnMeanDist* and *LnStdDist*. The boxplot distribution of the values of the three variables are presented in Figure 11.

Table 2. ANOVA results – Performance variables.

|  |  |  |  |
| --- | --- | --- | --- |
| **Output**  **variable** | **LnMeanDist**  **p-value** | **LnsStdDist**  **p-value** | **T250**  **p-value** |
| P factor (pilot) | 3.3E-11 | 1.5E-11 | 6.0E-11 |
| S factor (scenario) | 0.068 | 0.24 | 0.72 |
| Interaction of factors P and S | 4.4E-03 | 4.6E-03 | 0.29 |

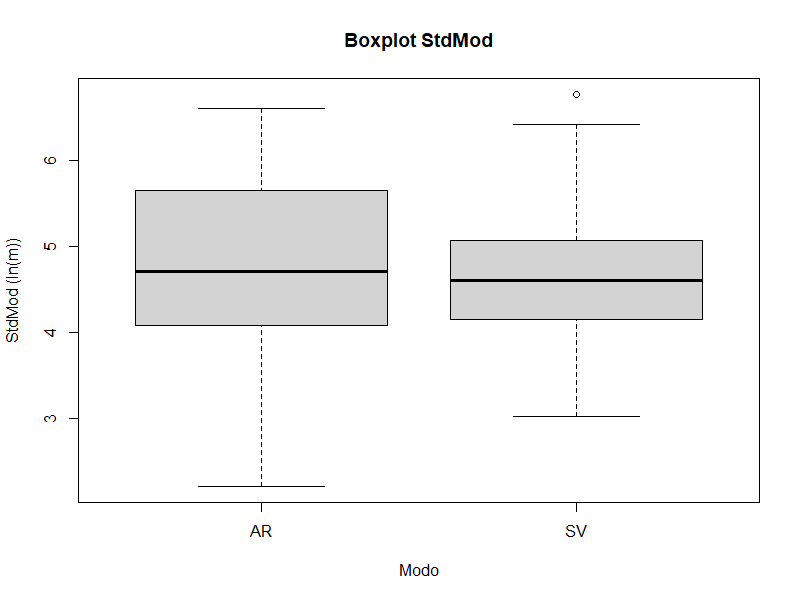
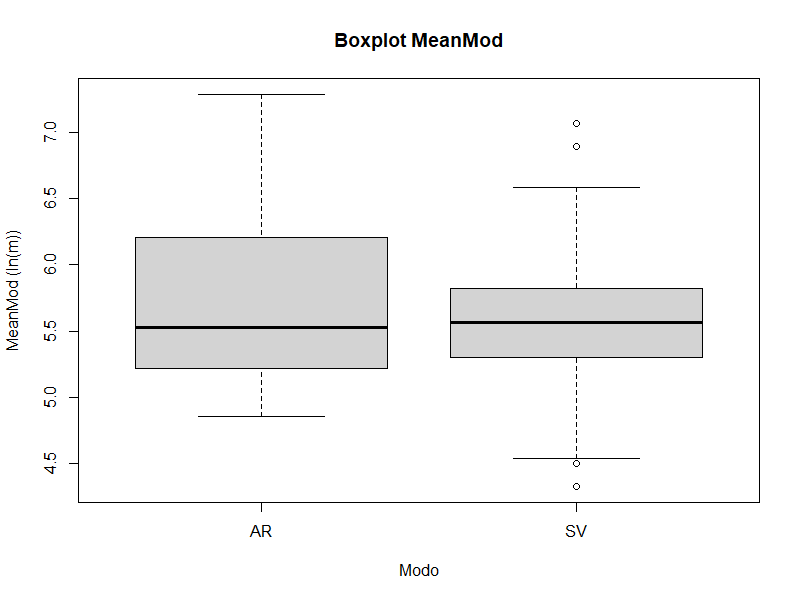
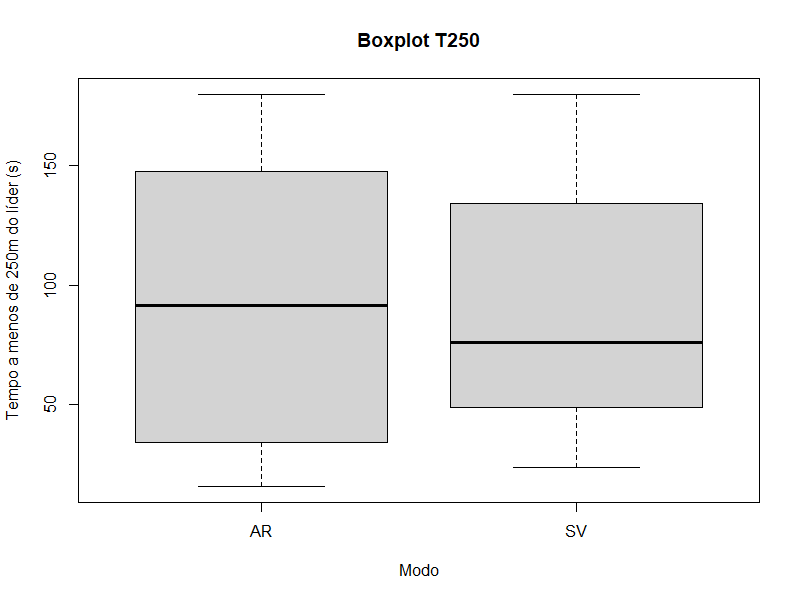
  

Figure 11. Boxplot of performance variables.

Summarizing the performance results, the hypothesis H1 “the pilot performance does not depend on the training scenario” could not be confirmed. On the contrary, the results showed that at least one performance variable (the distance between the airplanes) are influenced by the training scenario.

## Training effectiveness analysis – Hypothesis H2

Differently from the performance evaluation, each pilot trained in only one of the training scenarios. For the training effectiveness analysis, the variables *MeanDist*, *StdDist*, and *T250* were evaluated at two 1-min intervals: INI (at the beginning of the training session) and END (at the end of the training session). As a result, 20 values were initially obtained for each variable, corresponding to 10 for each scenario (AR or VS) and also 10 for each time (INI or END). The results of AR scenario were grouped in one dataset (AR dataset), while those of VS scenario were grouped another (VS dataset). Each dataset was analysed individually to check whether or not the performance at the INI interval is significantly different from the performance at the END interval.

Similar to what was done in the previous session, the variables were submitted to the Barlett and Shapiro Wilk tests, and obtained p-value was above the acceptance limit (p-value > 0.1). Following, ANOVA was applied using the statistical model of Eq. 5. In this case, there is no repetition for a given pilot and training scenario, therefore the pilot factor is blocked:

(5)

where:

* : output variable (*MeanDist*, *StdDist*, and *T250*);
* : average of the output variable;
* : variance due to T factor, related to timing in the training session (INI or END);
* : blocked influence of the P factor;
* : random error.

The results are presented in Table 3 and shows that in general the timing has a significant influence on the performance variables, i.e., the pilots’ performance is affected by training session - the only exception is for the *StdDist* variable of the AR dataset. As expected, for most of the pilots, the mean distance between the airplanes and its standard deviation were significantly reduced after training, while the amount time the distance was within the limit of 250 m was increased.

Table 3. ANOVA results – Training effectiveness variables – T factor.

|  |  |  |  |
| --- | --- | --- | --- |
| **Output**  **variable** | **MeanDist**  **p-value** | **StdDist**  **p-value** | **T250**  **p-value** |
| T factor (timing) AR dataset | 0.023 | 0.15 | 0.043 |
| T factor (timing) VS dataset | 0.093 | 0.046 | 0.050 |

Following, the variables *DiffMeanDist*, *DiffStdDist*, and *DiffT250* were calculated as the difference between the corresponding variables INI and END. They are used to estimate how the pilot performance was affected by the training session. As a result, 10 values were obtained for each variable, corresponding to 5 for each scenario. This dataset was then submitted to the Shapiro Wilk and Barlett tests and analysed using a model similar to that of Eq. 5 but, in this case, the T factor was replaced by the S factor, with the purpose of comparing the improvement in performance obtained with each training scenario. The results are presented in Table 4 and indicates that none of the variables are affect by the training scenario, i.e., the improvement in performance obtained with the using the augmented reality solution is not significantly different from that obtained without it.

Table 4. ANOVA results – Training effectiveness variables – S factor.

|  |  |  |  |
| --- | --- | --- | --- |
| **Output**  **variable** | **MeanDist**  **p-value** | **StdDist**  **p-value** | **T250**  **p-value** |
| S factor (training scenario) | 0.56 | 0.95 | 0.71 |

Summarizing the analysis of the training effectiveness, the hypothesis H2 “the training effectiveness does not depend on the training scenario” is considered confirmed by the experiment. The results confirms that the pilot performance is improved during the training session and there is no significant difference between the improvement obtained in each training scenario.

## Workload and stress analysis – Hypothesis H3

The analysis of hypothesis H3 is composed of the analysis of NASA-TLX questionnaire and the analysis of physiological data.

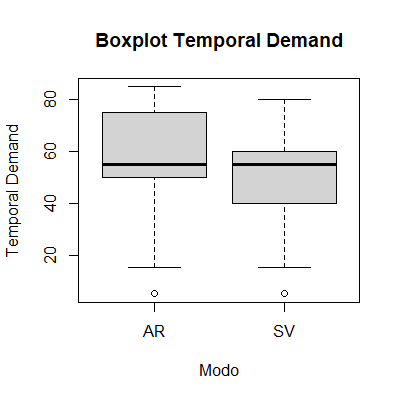
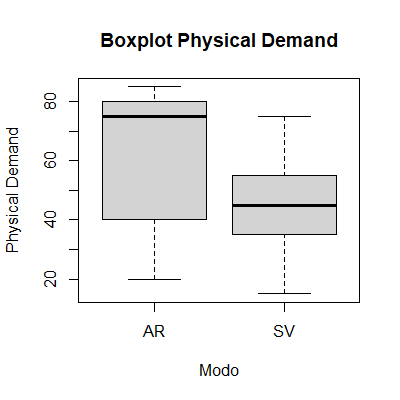
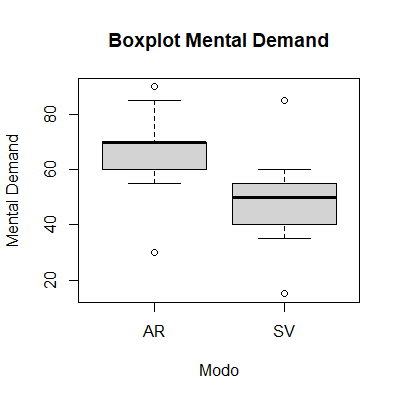
The NASA-TLX questionnaire was answered twice by each pilot, after the evaluation step, one corresponding to each training scenario. As a result, for each output variable, a total of 20 values should have been obtained. However, one of the pilots misunderstood the questionnaire and his answers were eliminated from the analysis, reducing the dataset to a total of 18 values for each variable.

Following the sequence of analysis adopted in the previous sections, the data were submitted to the Barlett and Shapiro Wilk tests and the results obtained were above the minimum limit (p>0.1). They were then analysed considered a model with one single factor (S factor, corresponding to the training scenario), similar to that of Eq. 5. The p-value obtained from the ANOVA analysis is presented in Table 5.

Table 5. ANOVA results – NASA TLX variables.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Output**  **variable** | **Mental demand**  **p-value** | **Physical demand**  **p-value** | **Temporal demand**  **p-value** | **Performance**  **p-value** | **Effort**  **p-value** | **Frustration**  **p-value** | **Weighted mean**  **p-value** |
| S factor (training scenario) | 0.052 | 0.028 | 0.41 | 0.031 | 0.030 | 0.010 | 0.036 |

Differently from what was initially expected, almost all the NASA-TLX variables indicated that the training scenario has a significant influence in the pilots’ perceived workload. The corresponding boxplot are presented in Figure 12 and Figure 13. They clearly indicate an increase in the values attributed to each NASA-TLX dimension for the training scenario where the pilot uses the augmented reality system.



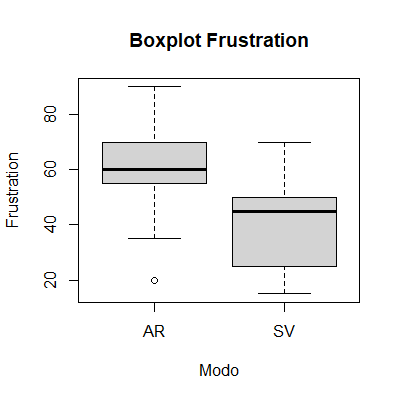
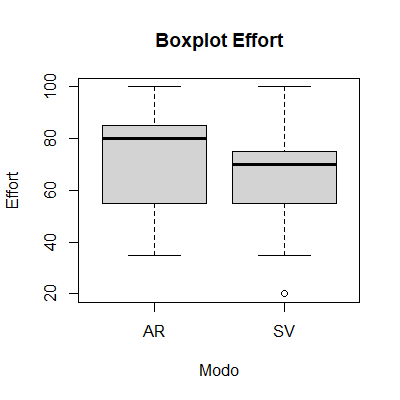
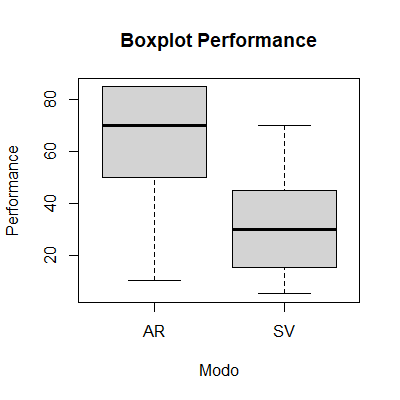


Figure 12. Boxplot of NASA-TLX factors.

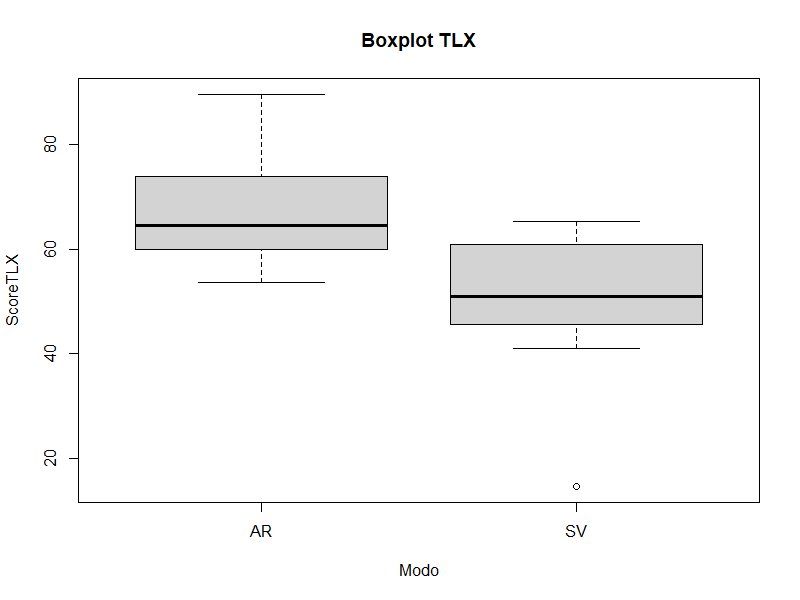


Figure 13. Boxplot of NASA-TLX weighted mean.

We conjecture that the increase in the perceived workload may be due to the fact that it can be easier to lose sight of the leader aircraft when using the augmented reality solution, due to the limited field of view of the glasses. Also, the difference in the brightness and sharpness of the projection may require more effort and attention from the pilot.

Following, we proceed to the analysis of the data from the physiological sensors. The data obtained from the ECG and EDA sensors were used to determine the following variables: *BPM*, *RRi*, *LF/HF* and *EDA*. For each pilot and each variable, a total of 11 values are obtained: 6 values for each 3-min flying interval (3 for each scenario) and 5 values from the intermediate relaxing 1-min intervals.

The statistical model considered for the ANOVA is similar to that of Eq. 4 and considers two factors: P for the pilot and I for the type of interval. The I factor can assume three different levels: AR, which corresponds to interval where the pilot is flying with the augmented reality solution, SV, which corresponds to interval where the pilot is flying without the augmented reality solution, and LX, which corresponds to a relaxing interval. As for the previous analysis, the Shapiro Wilk and Barlett tests were applied. The obtained p-values confirmed the normality of the data distribution and the homogeneity of the variances, with the exception of that corresponding to the Barlett test for the P factor of the EDA variable (p-value = 0.006). As the hypothesis of homogeneity is not confirmed, the ANOVA results of the EDA variable were excluded from the analysis.

The ANOVA results for the remaining variables are presented in Table 5. It shows a strong influence of the pilot factor in all the four output variables. On the other hand, none of the variables were sensitive to the type of interval, while the interaction of factors where relevant for the two time-domain variables of the ECG.

Table 5. ANOVA results – ECG and EDA variables.

|  |  |  |  |
| --- | --- | --- | --- |
| **Output**  **variable** | **BPM**  **p-value** | **RRi**  **p-value** | **LF/HF**  **p-value** |
| P factor (pilot) | ≈0 | ≈0 | 2,8E-09 |
| I factor (type of interval) | 0.14 | 0.16 | 0.41 |
| Interaction of factors P and I | 0.015 | 0.020 | 0.61 |

Summarizing the analysis, hypothesis H3 “the pilot workload and stress do not depend on the training scenario” is rejected based on the results obtained from the NASA-TLX questionnaire, which indicates an increase in the workload perceived by the pilot. However, the workload increase was not sufficient to affect physiological measures. The ANOVA analysis of the ECG variables indicates no significant difference among the three types of intervals: training with augmented reality, training without augmented reality and relaxing interval. As the ECG variables are associated with both workload and stress, we considered this late result as an indication that the level of stress is not significantly different between the training scenarios.

# Conclusions

This work proposes the use of augmented reality for creating training scenarios in a real flight, where elements that are external to the aircraft can be visualized through augmented reality glasses. The proposal is evaluated using the case of training formation flights. The pilot under training is the wingman and the leader aircraft is the external element, created as a projection of the augmented reality glasses. In this case, the use of an augmented reality device is expected to provide cost reduction and increase safety, since the risk of in-flight collision is eliminated.

As a necessary step in the development of this training system, a set of experiments was carried out in a simulation environment, where the formation flight training using augmented reality is compared with the case of the leader aircraft projected by the visual system of the flight simulator. A prototype was developed and integrated into a high-fidelity flight simulator available at the ITA. The prototype is composed of three systems: (a) a plug-in of the simulator's visual system; (b) an augmented reality system, based on the use of Hololens glasses, with wi-fi communication for data distribution; and (c) a central supervisory system, responsible for the mathematical calculations and generation of the trajectory, thus reducing the processing load to be embedded in the glasses. During the development of this prototype, different solutions were evaluated for referencing the pilot's position in relation to the aircraft's cockpit, including standard positioning methods, recognition using markers and image-based recognition. The adopted solution combines standard positioning with a manual calibration system, where the pilot adjusts the position of a mask (holographic image of the cockpit) to superimpose it on the real cockpit of the flight simulator.

In order to evaluate the impact of the augmented reality system on the pilot, different factors are analysed from the data collected in an experimental training campaign, including flight data, as well as data obtained from physiological sensors and subjective questionnaires. Three hypotheses are defined, corresponding to the absence of influence of the augmented reality system on the following factors: performance, training effectiveness, and workload and stress.

The experiment is composed of two steps. The first step provides an indicative line of the evolution of the pilot's performance throughout a training session, considering that each pilot uses only one training environment configuration (with or without augmented reality). The second step includes the comparison, for each pilot, of the two configurations, through alternating training in both configurations, interspersed with relax intervals.

The experiment was carried out with a group of 10 volunteers. The results obtained from the analysis of the collected data confirmed the hypotheses that the training effectiveness is not affected by the use of augmented reality. On the other hand, the other two hypothesis were rejected. Concerning performance, the ability to maintain a certain distance from the leader is affected by the use of augmented reality glasses. Additionally, the results obtained from the NASA-TLX questionnaire show a clear increase in the workload with the use of the augmented reality glasses.

As future work, we plan to adapt the prototype to new augmented reality glasses, such as Hololens 2, which presents a large field of view, processing power and positioning accuracy. We also plan to test in-house algorithms for image positioning, based on data fusion from different sensors, including, but not limited to, the ones embedded in the glasses, with the purpose of enabling its use in dynamic environments.

## Acknowledgements

This work was supported by CAPES, FAPESP (Process 2012/51085-3), FINEP/VINNOVA (Process 01.17.0038.00), CNPq (Process 303271/2017-5).

# References

1. ALEXANDER, T. et al. Enhancing Human Effectiveness through Embedded Virtual Simulation. I/ITSEC Interservice/Industry Training, Simulation, and Education Conference, 2012.
2. ARTHUR, J. J. et al. A review of head-worn display research at NASA Langley Research Center. SPIE Defense + Security. Baltimore: SPIE, 2015.
3. ARTHUR, J. J. et al. A review of head-worn display for the Next Generation Air Transportation System. SPIE Optical Engineering, Baltimore, v. 56, 2017.
4. AUER, S. et al. Comparison Between Virtual Reality and Physical Flight Simulators for Cockpit Familiarization. MundC: Mensch und Computer, 2021.
5. BCAA. Aviation Safety Information Leaflet Formation Flying. Belgian Civil Aviation Authority, 2016.
6. BELL, H. H.; WAAG, W. L. Evaluating the Effectiveness of Flight Simulators for Training Combat Skills: A Review. THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, v. 8, n. 3, p. 223 - 242, 1998. https://doi.org/10.1207/s15327108ijap0803\_4.
7. BLUME, B. D. et al. Transfer of Training: A Meta-Analytic Review. Journal of Management, 2010.
8. BOUCSEIN, W. Electrodermal Activity. 2ª. ed. New York: Springer, v. 1, 2012.
9. BRAITHWAITE, J. J. et al. A Guide for Analysing Electrodermal Activity (EDA) & Skin Conductance Responses (SCRs) for Psychological Experiments. University of Birmingham. Birmingham, p. 43. 2015.
10. BÜRKI-COHEN, J.; GO, T. H. The Effect of Simulator Motion Cues on Initial Training of Airline Pilots. AIAA Modeling and Simulation Technologies Conference and Exhibit. San Francisco: AIAA. 2005. p. 12.
11. BÜRKI-COHEN, J.; SPARKO, A. L.; BELLMAN, M. Flight Simulator Motion Literature Pertinent to Airline-Pilot Recurrent Training and Evaluation. AIAA Modeling and Simulation Technologies Conference. Portland: AIAA. 2011. p. 17.
12. DICKSTEIN, C. Air Force instructor, student pilot killed in T-38 crash at Vance Air Force Base. Stars and Stripes, 2019.
13. FAA. "Flight Simulation Training Device Initial and Continuing Qualification and Use," 14 CFR Part 60, NSP Consolidate Version 2016.
14. FOOTE, B.; MELZER, J. A history of helmet mounted displays. SPIE Defense + Security. Baltimore: SPIE. 2015.
15. GAWRON, V. J. Human Performance, Workload, and Situational Awareness Measures Handbook, 2008.
16. HART, S. G. Nasa-Task Load Index (NASA-TLX); 20 Years Later. Proceedings of the Human Factors and Ergonomics Society 50th Annual Meeting, 2006. p. 904-908.
17. HART, S. G.; STAVELAND, L. E. Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Human Mental Workload, 1988.
18. HAYS, R. T. et al. Flight Simulator Training Effectiveness: A Meta-Analysis. Military Psychology, v. 4, p. 63-74, 1992.
19. HIDALGO-MUÑOZ, A. R. et al. Cardiovascular correlates of emotional state, cognitive workload and time-on-task effect during a realistic flight simulation, 2018.
20. JOHANNES, B. et al. Psychophysiological Assessment in Pilots Performing Challenging Simulated and Real Flight Maneuvers. Aerospace Medicine and Human Performance, v. 88, n. 9, p. 834 - 840, 2017.
21. KRIJN, R.; WEDZINGA, G. Development and in-flight demonstration of 'E-CATS', an experimental embedded training system for fighter aircraft. National Aerospace Laboratory NLR. p. 25. 2004. (NLR-TP-2004-393).
22. LABEDAN, P. et al. Virtual Reality for Pilot Training: Study of Cardiac Activity. n Proceedings of the 16th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications, 2021.
23. LEE, A. T. Flight Simulation: Virtual Environments in Aviation. Nova York: Taylor & Francis Group, v. 1, 2016.
24. LEINO, T. Neuroendocrine Responses to Psychological Workload of Military Flying. Oulo: Oulu University Library, 1999.
25. LEMMERS, A. J. J. An Embedded Training Multi-Ship Demonstrator. National Aerospace Laboratory NLR. p. 17. 2008. (NLR-TP-2008-005).
26. MACFARLANE, P. W. et al. Comprehensive Electrocardiology, 2011.
27. MAGEE, L.; SOTTILARE, R.; ROESSINGH, J. J. Human Interaction in Embedded Virtual Simulations. ITEC. Cologne: ITEC. 2011.
28. MANSIKKA, et al. Fighter pilots' heart rate, heart rate variation and performance during an instrument flight rules proficiency test. Applied Ergonomics, 2016.
29. MARINESCU, A. et al. Physiological Parameter Response to Variation of Mental Workload. Human Factors, p. 31 - 56, 2018.
30. MCCLERNON, C. K. Stress Training Efficacy in an Aviation Context. International Symposium on Aviation Psychology, 2009.
31. MCCLERNON, C. K. et al. Stress Training Improves Performance During a Stressful Flight. Human Factors, v. 53, n. 3, p. 207-218, 2011.
32. MCSPADDEN, R. What Went Wrong? Formation Flight Turns Tragic. AOPA, 2019.
33. ORITZ, G. A. Transfer of Learning Effectiveness: PC-Based Flight Simulation. Journal of Aviation/Aerospace Education & Research, v. 3, n. 2, p. 29-33, 1993.
34. REID, G. B. Training Transfer of a Formation Flight Trainer. Human Factors, 1975. 470-476.
35. ROBINSON, A. D.; BICANIC, N. Methods and Systems to Allow Real Pilots in Real Aircraft Using Augmented and Virtual Reality to Meet in a Virtual Piece of Airspace. US 2019/0333396, 2019.
36. ROESSINGH, J. J. M. Transfer of Manual Flying Skills From PC-Based Simulation to Actual Flight—Comparison of In-Flight Measured Data and Instructor Ratings. THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, v. 15 (1), p. 67 - 90, 2005.
37. ROESSINGH, J. J. M.; VAN SIJLL, M. C.; JOHNSON, S. P. Embedded Training - An explorative study providing requirements for the display of virtual targets on a Helmet Mounted Display in simulated air-to-air engagements within visual range, Amsterdam, p. 54, 2003. ISSN NLR-TP-2003-262.
38. ROESSINGH, J. J. M.; VERHAAF, G. G. Training Effectiveness of Embedded Training in a (Multi-) Fighter Environment A Discussion Paper. National Aerospace Laboratory NLR. p. 8. 2009.
39. SHAFFER, F.; GINSBERG, J. P. An Overview of Heart Rate Variability Metrics and Norms. Frontiers in Public Health, vol. 28, 2017.
40. TASK GROUP HFM-165. Improving Human Effectiveness Through Embedded Virtual Simulation. OTAN, 2014.
41. VALVERDE, H. H. A Review of Flight Simulator Transfer of Training Studies. Human Factors, Ohio, v. 15, n. 6, p. 512 - 522, 1973. ISSN https://doi.org/10.1177/001872087301500603.
42. VELTMAN, J. A. Comparative Study of Psychophysiological Reactions During Simulator and Real Flight. The International Journal of Aviation Psychology, v. 12, p. 33 - 48, 2002.
43. WEDZINGA, G. E-CATS: First time demonstration of embedded training in a combat aircraft. Aerospace Science and Technology, v. 10, p. 73-84, 2005.
44. WOOD, M. E. et al. Design of the Simplified Formation Trainer. Defense Technical Information Center. Ohio, p. 36. 1972. (AD0754973).