



HAL
open science

Comparison and validation of an immersive flight simulator using a virtual reality headset versus a professional simulator for novice pilots

Johan Rendy, Isabelle Milleville-Pennel, Cédric Dumas, Thierry Lebourque

► To cite this version:

Johan Rendy, Isabelle Milleville-Pennel, Cédric Dumas, Thierry Lebourque. Comparison and validation of an immersive flight simulator using a virtual reality headset versus a professional simulator for novice pilots. *Ergonomics*, 2025, pp.1-19. <10.1080/00140139.2025.2525956>. <hal-05381423>

HAL Id: hal-05381423

<https://hal.science/hal-05381423v1>

Submitted on 27 Nov 2025

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Copyright - All rights reserved

1
2
3
4
5
6 **Comparison and validation of an immersive flight simulator using a virtual**
7 **reality headset versus a professional simulator for novice pilots**
8
9

10
11 Johan Rendy ^{ad}, Isabelle Milleville-Pennel^{abc*}, Cédric Dumas^{ea} & Thierry
12 Lebourque^d
13

14
15 *^aLS2N-UMR CNRS 6004, Nantes, France; ^bNantes University, Nantes, France; ^cEcole*
16 *Centrale de Nantes, Nantes, France; ^dAlsim, Le Loroux-Bottereau 44430, France;*
17 *^eIMT Atlantique Bretagne-Pays de Loire DAPI, Nantes, France;*
18
19

20
21 *Corresponding Author: isabelle.milleville@ls2n.fr; Orcid: [https://orcid.org/0000-0002-](https://orcid.org/0000-0002-2236-8346)
22 [2236-8346](https://orcid.org/0000-0002-2236-8346)
23
24

25
26
27
28 Runnig Head: Validation of headset flight simulator
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Abstract

The aim of this study was to assess whether a low-cost flight simulator, using a virtual reality (VR) headset and specially adapted to the needs of novice pilots, could meet training expectations. This simulator was compared with a traditional professional flight simulator equipped with a CAVE projection room. The participants underwent flight training in one or other of the simulators and their performance was compared. Subjective evaluations showed high level of presence despite some dimensions about the feeling of presence sometimes indicates a difference in disfavour of the Headset simulator. Similar mental workload among participants was also observed as well as nearly comparable learning levels (flight performances) between the two systems.

Our study demonstrated that a Headset simulator can provide an immersive and educational experience comparable to that of a Cockpit simulator. Thus, Headset simulator represent a viable and economically advantageous alternative to traditional Cockpit simulators for initial pilot training.

Keywords

Flight simulator, pilot training, virtual reality, presence, mental workload

Practitioner Summary

In this paper, it is experimentally shown that a low-cost and functionally simplified flight Headset simulator can provide an immersive and educational experience, on a specific task, comparable to that of a professional Cockpit simulator. These results underscore the importance of continuing to develop affordable training solutions without compromising the quality of learning.

Introduction

Objectives

Simulators are now widely used in the field of training, and the aeronautics sector is no exception. Numerous studies have shown that training with a simulator can be transferred to a real flight situation and can be beneficial for learning (Koglbauer et al, 2016; Macchiarella, 2006; Marques et al, 2022). In some cases, it can even shorten the learning process (Marques et al, 2022). Today, there are several types of professional simulator, as listed by the European Aviation Safety Agency (EASA). Each of them meets specific training needs.

However, there are currently no simplified, "low-cost" simulators for aeronautical training, even though there is a real need for them. Indeed, the first stages of learning to fly do not require the use of all flight controllers and rely on very specific and limited procedures and adapted learning methods (Marques et al, 2022, Rendy, 2023). From this point of view, a simplified, lower-fidelity, simulator could be appropriate, provided it delivers sufficient information to enable good situational awareness and learning of the main actuators.

Existing technologies

Recent years have seen the emergence of high-performance VR headsets, making VR more accessible. As a result, this technology is developing in the field of flight simulation, particularly in the public and commercial markets with video game licenses such as XPlane 11 by Laminar Research (2017) or Flight Simulator by Asobo Studio (2020). The latter uses headsets with a fully virtual simulation featuring visuals of the cockpit and the entire flight environment. This makes it easier to reconfigure the aircraft used, the training subject or even the aircraft's position. Little or no physical hardware is used, and vision is almost the only sense involved. This type of device is less expensive and takes up less space. In some cases, it is used in aviation schools as a complementary training tool. However, as it does not allow tactile and haptic interaction, it is considered as having a low fidelity. So, it is therefore limited in its contribution to training. However, Paul et al. (2025), showed the potential benefits of implementing VR simulators in ab initio pilot training. In their study, VR simulation was found to be effective at training procedural tasks with a high level of cognitive fidelity. In particular, they provide situational awareness equivalent to that of an ALSIM AL250. However, the authors also showed that haptic feedback from flight controls is necessary for better learning. To achieve higher fidelity, other players in the field are attempting to combine VR headsets with professional simulators, such as the simulator by Oberhauser et al. (2018). This simulator materially and computationally reproduces the entire cockpit and behavior of an aircraft while using a VR headset to immerse the user in a totally virtual environment. However, these devices offer few advantages over a traditional simulator, since they still have a cockpit that exactly reproduces the physical system existing in reality, to which is added a VR headset for visuals, instead of the projection screens usually used. Although it offers high fidelity, it is expensive and takes up a lot of space. In addition, it requires a precise configuration of the aircraft corresponding to the tangible support of the cockpit.

Our prototype

Both types of simulators have their advantages. The first, using VR headsets only, despite offering less fidelity, is much less expensive, less bulky, transportable and reconfigurable. The second type, can be used to appeal to different senses, providing higher fidelity with more sensory stimuli - an essential point for training. In order to propose an optimal simulator, a prototype was therefore developed that takes advantage of the benefits of both types of simulators. To achieve this, the latest VR headset products available on the market were used to propose a fully immersive, compact and simplified simulator prototype, which retains most

1
2
3
4 of the tangible actuators and sensory information useful in the very early stages of learning.
5 According to Macchiarella et al (2006), the use of a flight training device can be useful for
6 learning, provided that the strengths and weaknesses of the device are taken into account when
7 defining the nature and/or the part of the activity to be learned. The work of Rendy (2023) was
8 therefore used to identify the information essential to the pilot at the very start of learning and
9 to propose a simplified version of the actuators which, despite lower fidelity, preserves these
10 functions (see section *materiel and methodology* for more details).

11 It therefore remains to be demonstrated that a headset simulator, specially designed for
12 beginner novice pilots, can be as effective as a standard Cockpit simulator not optimized for
13 this task. To this end, the two types of flight simulators were compared when training novice
14 pilots. More specifically, this work will consider the influence of visual immersion technology
15 and haptic feedback on the subjective feeling and performance of novice pilots when learning
16 a specific task during the very first steps of confrontation to virtual flying.

17 *Prerequisites for validation*

18 As far as subjective experience is concerned, two aspects are important to consider in order to
19 validate psychological fidelity of the simulator: the feeling of presence and the mental load felt.
20 The feeling of presence can be defined as a binary experience, in which the perception of self-
21 location and action possibilities is linked to a mediated spatial environment. Similarly, the
22 mental functions called upon are linked to this environment (Wirth et al., 2007). Thus, the
23 feeling of presence is multidimensional and must be questioned as such. It is essential that the
24 headset simulator achieves at least the same level of presence as the Cockpit simulator, in the
25 different dimensions of presence.

26 Concerning the Mental workload (MWL), it represents the degree of activation of a finite pool
27 of resources, limited in capacity, while cognitively processing a primary task over time,
28 mediated by external stochastic environmental and situational factors, as well as affected by
29 definite internal characteristics of a human operator, for coping with static task demands, by
30 devoted effort and attention (Longo et al., 2022). To enable repeated use of the simulator, it is
31 essential to limit the MWL associated with its use, or at least to ensure that it is equivalent from
32 one simulation technology to another. MWL can be questioned through several dimensions
33 such as those considered in NASA-TLX (Cegarra & Morgado, 2009) and directly linked to the
34 piloting task, such as mental load, stress or the effort felt to perform the task. This questionnaire
35 is commonly used in aviation to measure pilots' mental workload (Mansikka et al, 2019; Wright
36 & O'Hare, 2015; Zheng et al, 2019). Dimensions more closely linked to the perceptual cost of
37 the piloting task can also be considered, such as visual processing or visuo-spatial processing,
38 investigated by Workload Profile (WP) test (Tsang & Velazquez, 1996). It has been
39 demonstrated that WP has a high diagnostic power (Valdehlta et al. 2004) and it advisedly
40 complements Nasa-TLX measure.

41 Furthermore, as the main objective of the simulator is training, it must be possible to
42 demonstrate its performance in terms of learning. To do this, it is important not only to measure
43 changes in flight performance during simulator training, but also to assess their quality by
44 comparing them with those of an expert pilot on the two training devices.

45 Finally, it is necessary to ensure that the headset simulator will not cause more discomfort to
46 users than the Cockpit simulator. Such as simulator sickness, assessed using the Simulator
47 Sickness Questionary (SSQ) developed by Kennedy (1993) and translated into French by
48 Bouchard et al. (2009).

49 *Objective and Hypothesis*

50 Our aim in this study is therefore to show that a simplified flight simulator, with lower haptic
51 fidelity, can achieve the same level of learning and performance during a simulated flight as a
52
53
54
55
56
57
58
59
60

1
2
3
4 more powerful cabin simulator normally used in aviation schools, as far as the very early stages
5 of learning are concerned.

6 We therefore hypothesis that user performance, the feeling of presence and the mental load will
7 be equivalent between the two simulators, and that a headset simulator does not induce a greater
8 risk of simulator sickness than a cabin simulator.
9

10 11 Materiel and Methodology

12 13 *Experimental setup*

- 14 • Alsim Cockpit simulator AL250

15 The first simulator, referred to as the 'Cockpit' simulator, is the AL250 professional Cockpit
16 simulator marketed by Alsim (fig. 1a and table 1). The AL250 simulator is certified by the
17 EASA (European Union Aviation Safety Agency) for training purposes. It corresponds to the
18 types of models encountered in many aviation schools. The fidelity of this simulator is an
19 approved flight training device (FTD) so that a certain number of flight hours are authorized
20 on this simulator in place of real flights. It does not replicate any specific type of existing
21 cockpit, but it is reconfigurable to simulate a range of aircraft. The Cockpit simulator is
22 equipped with a motorized force feedback steering column, which allows for immersion and
23 control identical to that of a real aircraft. Indeed, the Cockpit simulator is certified by aviation
24 authorities, ensuring that the behavior of the steering column (and the simulator as a whole)
25 matches that of a real aircraft in the same category. The Cockpit simulator features three
26 projectors and three flat projection surfaces. However, a limitation is noted when the pilot turns
27 their head, exposing the screen supports, which affects the visual coverage.
28
29

30 / Insert figure 1 about here/
31

32 / Insert table 1 about here/
33

- 34 • Experimental Headset simulator

35 Our second device is an experimental simulator specifically crafted for this study. The Headset
36 simulator is a single-seat flight simulator that uses a virtual reality headset (fig. 1b and table
37 1). It comprises a physical part and a virtual part. The physical part consists of a steering
38 column (table 2), a throttle lever, and rudder pedals that have 'passive' haptic feedback,
39 meaning they are non-motorized and operate via springs that offer little resistance
40 comparatively to the other system (where up to 31 times more resistance is observed).
41

42 / Insert table 2 about here/
43

44 To ensure alignment with the virtual environment, yoke is not directly tracked; instead, its data
45 are transmitted via a USB signal. Leap Motion is used to track hand movements. The HTC
46 Vive Pro Eye headset is tracked using HTC base stations. The simulator's buttons are manually
47 reset to their initial positions before the system is started. Audio is delivered through a stereo
48 headset. The visuals are provided by the HTC Vive Pro Eyes headset. The virtual component
49 consists of an external representation of the airplane, displaying the flight environment in the
50 same manner as the AL250 (same visual environment outside the cockpit). The headset
51 simulator features an airplane inside cockpit developed specifically for this simulator, along
52 with a dashboard, onboard instruments, and virtual control devices (fig. 1b bottom). In
53 developing the dashboard and its buttons for the VR simulator, we relied on preliminary
54 findings focused on flight activity model (Rendy, 2023). This work played a crucial role in
55 defining the ergonomic and functional aspects of our simulation interface. It demonstrated the
56
57
58
59
60

1
2
3
4 importance of haptic feedback depending on the controls. Indeed, for some buttons, precise
5 haptic feedback is not essential, unlike flight controls where feedback is crucial. So, the
6 Headset simulator does not feature physical buttons on its dashboard unlike the Cockpit
7 simulator, instead a simplest passive-haptic feedback technology was chosen (when pushing
8 on a button, the participant feels the contact with the simulator board to which the steering
9 column is fixed). This simplify interaction has been showed as promising in many interaction
10 tasks (Kim, Kim & Jo, 2023). As far as visualization in the virtual environment is concerned,
11 the user is represented by a virtual view of his/her hands. Furthermore, Rendy (2023)
12 highlighted the varying importance of controls, emphasizing the need to include only essential
13 controls for the first piloting stages.

14 So that we can then correctly identify the benefits and limitations in terms of learning, it is
15 essential to recognize the inherent limitations of each technology, Cockpit and Headset
16 simulator. The Cockpit simulator, despite its accredited realism, has visual constraints that
17 might limit immersion. Conversely, the VR simulator, while providing unprecedented visual
18 immersion, lacks realistic haptic feedback, which could affect the learning of the finer nuances
19 of flight control.
20

21 22 23 *Participant*

24 The experiment involved 22 volunteers (12 men, 10 women) with an average age of 24.9 years
25 (ranging from 19 to 53 years) who were divided into 2 groups. Each participant was
26 compensated 100€. Group 1 (Headset simulator) consisted of 12 participants (7 men and 5
27 women) and Group 2 (Cockpit simulator) of 10 participants (5 men and 5 women). They were
28 chosen for having no prior flight training or flight simulator experience. They had no vision
29 impairments or were corrected to normal through surgery or lenses so that they could wear
30 them with the headset. Participants were recruited from the University of Nantes, the École
31 Centrale de Nantes, and the social circles of the doctoral student.

32 A pilot who was also an instructor participated in designing the tutorial for the presentation and
33 training session of the flight. This 44-year-old expert holds CPL(A), IR-ME, IR-SE, ULM, and
34 FI(A) certifications, including night flying. This instructor was employed by the company
35 involved in developing the simulator. He was piloting different sort of planes since 2002 and
36 had 580 of flying hours on the sort of plane considered in the study (Alsim single-engine
37 propeller).
38

39 40 *Task definition*

41 The flight stage selected for the study was the left-hand circuits. A 'circuit' in aviation is a
42 standard procedure performed by an aircraft when it is around the airport. It includes several
43 stages that will be describe in the next section. This procedure ensures a safe and orderly flow
44 of air traffic, particularly for light and training aircraft, by following predefined trajectories and
45 allowing for a quick response in case of a go-around for safety reasons. Since the aim of this
46 work was to study individuals with no prior flight knowledge, the task was simplified with
47 respect to piloting by restricting it to the manipulation of the aircraft alone. Thus, the parts
48 involving taxiing to the runway, checklist, returning to parking, and communication with the
49 control tower were omitted. Apart from the information given by the flight instruments (which
50 was identical for both simulators), there was no additional feedback on performance (ie. no
51 oral or qualitative feedback on flight performance).
52

53 54 *Flight Scenario*

55 During the scenario, instructions are given to the participant based on Aeronautical headings.
56 These indications are tools for navigation that allow airplane pilots to know the orientation of
57
58
59
60

1
2
3
4 their aircraft relative to magnetic north. Thus, navigation tools measure the angle between the
5 direction in which the aircraft is moving and magnetic north and display this information on a
6 dial graduated in degrees. In our flight scenario, 4 directions are important (25°, 115°, 205°,
7 and 295°) as they allow to localize 3 important visual landmarks corresponding to a forest, a
8 tennis court and the runway (fig. 2). For the experiment, the participant virtually located at the
9 Cholet aerodrome.

10
11 / Insert figure 2 about here/
12

13 The figure 2 represents the instructions and areas of interest that the participants must follow :

14 - *Step 1 : Take off*

15 The first expected action was to take off; for this, the participant was to go full throttle and then
16 stabilize the aircraft on the center of the runway using the rudder pedals. Once a speed of 60
17 knots was reached and the aircraft was stabilized, the initial climb was initiated by pulling on
18 the stick.

19 - *Step 2 : Initial climb*

20 During the initial climb phase, the participant had to keep the throttle at maximum and maintain
21 the stability of the aircraft. The orientation of the runway at takeoff was 25, so the participant
22 had to maintain this heading. He/she also has to maintain the climb angle indicated during
23 training, known as the climb attitude. Once the aircraft has climbed 500 feet, corresponding
24 here to an altitude of 940 feet (since the aerodrome is at an altitude of 440 feet), the participant
25 was to continue climbing, but begin a 90° turn with a 20° bank to enter the crosswind leg.

26 - *Step 3 : Crosswind*

27 The participant continued to climb to 1440 feet (+1000 feet altitude relative to the runway
28 threshold) and kept turning to reach a heading of 295°. Then, the participant was to proceed
29 until the forest (fig. 2) was on their left side and initiate a 90° turn with a 20° bank to enter the
30 downwind leg.

31 - *Step 4 : Downwind*

32 After reaching an altitude of 1440 feet and stabilizing their trajectory, the participant was to
33 complete the turn by orienting the aircraft toward a heading of 205°, which made it parallel to
34 the runway. Then, they were to maintain a stable position with the wings level and parallel to
35 the runway until moving away from and passing it, so that (when turning their head 45° from
36 the rear) they could see the end of the runway (a top-down view of the situation can be seen in
37 Figure 3), while keeping the specified landmark in view (in this case, the tennis courts, fig. 2).
38 Upon reaching this landmark, the pilot was to execute a left turn of 90° with a 30° bank.
39
40

41
42 / Insert figure 3 about here/
43

44 - *Step 5 : Base*

45 The participant was to position the aircraft perpendicular to the runway with a heading of 115°.
46 Once out of the turn, he/she was to reduce the throttle (setting it between 1900 and 2000 rpm)
47 and begin the descent as instructed in the tutorial. Once nearly aligned with the runway and at
48 an altitude of about 940 feet, the participant was to turn left again by 90° and thus align with
49 the runway, with a 20° bank.

50 - *Step 6 : Final and landing*

51 The participant was to continue descending while aligning and centering on the runway. The
52 throttle was to be progressively reduced as the participant approached the runway, cutting it
53 off about 1 meter above the ground. Before touching down on the runway, the participant was
54 to slightly raise the nose of the aircraft to ensure the rear wheels touched down first.
55 Each participant faced the same flight complexities regardless of the simulator used.
56
57
58
59
60

Experimental Protocol

Participants were required to perform the “circuit” scenario on one of the two simulators, so two groups of participants were created, one for each simulator. For each group, the experiment was conducted over two learning sessions, each lasting about one hour. The learning sessions consisted of four stages for session 1 and three stages for session 2. The two sessions were spaced five to seven days apart.

Concerning session 1, the four stages were: “welcome and training”, “training flight”, “test flight” and “questionnaires completion”.

Concerning session 2, the three stages were: “training flight”, “test flight” and “questionnaires completion”.

- Welcome and Training :

Upon arrival, participants first filled out a consent and personal information document, then attended a training session during which they watched a video detailing the experiment, the basics of piloting, and the specific flight scenario. This video was developed in collaboration with the flight instructor (mentioned in the participants section) and ensured that all participants had a common base of information. At the end of the video, participants had the opportunity to review specific segments with the experimenter to clarify any points. Afterwards, additional information on visual landmarks and important aeronautical headings was provided, before the participants moved on to the training and testing stages on the simulator that had been assigned to them.

- Training Flight :

During session 1, this involved getting accustomed to the simulator for about 15 minutes, during which the experimenter answered questions, guided, and advised the participant so that he/she could apply what he/she had just seen in the video. Once the initial familiarization was completed, a one-minute break was taken, during which the participant remained seated in the pilot’s seat with their eyes closed, before continuing.

During session 2, it involved a refamiliarization with the simulator. This was done at the beginning of the session during 1 or 2 circuits, with a reminder of the important headings and visual landmarks. Once the refamiliarization was completed, a one-minute break, with eyes closed in the pilot's seat, was taken.

- Test Flight :

The participant took place in the simulator. Before starting, participants received the following instruction: 'If you realize that you are no longer piloting or that you have left the simulation, please let me know out loud by saying “lost”, to identify any lapses in Presence feeling.

Then, during session 1, the participant performed the flight scenario twice.

During session 2, the participant performed the flight scenario twice, took a one-minute break by closing his/her eyes, before performing the flight scenario twice again.

During all these circuits, data on position, angle, speed, and time were recorded. For the Headset simulator, a virtual scene recording system was used to capture what was happening in the virtual environment and was synchronized with an audio track recording the pilot. For the Cockpit simulator, the scene was recorded using a camera placed behind the passenger seat to record the dashboard and the external view.

- Questionnaires completion :

At the end of the last flight scenario of each session, participants were required to answer 5 questionnaires (see next section for details):

- The Presence Questionary, validated by Witmer et al. (1998) and translated into French by Robillard et al. (2002; QEP, for the french denomination “Questionnaire Etat de Presence”) for the presence.
- Pictorial Presence self-assessment manikins (Weibel 2015) for the presence.
- Nasa TLX for mental workload (Hart & Staveland 1988)
- WP for mental workload (Tsang 1996).
- Simulator Sickness Questionary for simulator sickness (Kennedy et al 1993).

At the end of each learning session, participants were invited to rest in the experimental room with the experimenter. They were only allowed to leave on their own if they showed no symptoms of simulator sickness or any other signs of fatigue or temporary discomfort. The experimental procedure is summarized in table 3.

/ Insert table 3 about here/

Measure

Four categories of measures were considered during each learning session : presence, mental workload, simulator sickness, and flight performances. For this purpose questionnaires were used and data from the simulator were recorded.

- *Presence measures* :

- Le Pictorial Presence self-assessment manikins (Weibel et al 2015) for the presence:

The SAM (Self-Assessment-Manikin) questionnaire was developed and validated by Weibel et al. (2015). It is a multidimensional questionnaire designed to subjectively measure user presence. This questionnaire is based on the MEC spatial presence model and the considerations of Wirth et al. (2007). To this end, the user positions themselves on different dimensions of presence represented by visual pictograms.

- Attention Allocation : This dimension measures the virtual environment's ability to attract and maintain the user's attention.
- Spatial Situation Model : This dimension measures the accuracy and clarity of the user's spatial representation of the virtual environment.
- Self- location: This dimension measures the extent to which the user feels located within the virtual environment rather than in reality.
- Possible Actions : This dimension measures the environment's ability to offer possibilities for actions and interactions to the user.
- Higher cognitive Involvement : This dimension measures the environment's ability to arouse the user's interest and cognitive engagement.
- Suspension of disbelief : This dimension assesses the user's willingness to suspend disbelief towards the environment and thus accept it as a possible reality.

The participant responds after the experiment by selecting the item on each line of pictograms that most corresponds to the situation, they have just experienced. Each scale was quoted from 1 to 5.

Presence is further assessed by a second presence questionnaire, described below.

- The QEP :

The Presence Questionary, validated by Witmer and Singer (1998) and translated into French by Robillard et al. (2002; QEP, for the french denomination “Questionnaire Etat de Presence”), complements the assessment of participants' subjective feeling of presence. The QEP consists

of 24 questions using a scale ranging from 0 to 100. This questionnaire is completed at the end of the session to avoid being intrusive.

For the analysis, the responses are grouped into seven dimensions :

- Realism subscale: This subscale assesses the extent to which the virtual environment seems realistic and credible.
- Possibility to act : This subscale assesses the extent to which the user feels they can act on the virtual environment and have an impact on it.
- Quality of interface : This subscale assesses the quality of the user interface in the virtual environment.
- Possibility to examine : This subscale assesses the user's ability to explore the virtual environment and discover new elements..
- Self-evaluation of performance : This subscale enables users to assess their own performance in the virtual environment.
- Sounds : This subscale assesses the quality of the sound effects and their realism in the virtual environment.
- Haptic : This subscale assesses the quality of haptic sensations and their realism in the virtual environment.

One of the advantages of this questionnaire over the SAM is that it allows for measuring presence in direct relation to haptic and auditory feedback, which was relevant in our study since the two simulators used different sensory rendering techniques. The QEP was thus complementary to the SAM. Together, these two questionnaires provide a comprehensive representation of presence in both simulators.

- *Mental workload measures :*

- The NASA TLX (Hart & Staveland 1988) for mental workload :

The questionnaire, developed by Hart and Staveland (1988), is recognized as being effective for measuring an individual's mental workload. It is a 6-question survey designed to measure six dimensions of mental workload :

- Mental Demand: This dimension refers to the amount of mental effort required to accomplish a task.
- Physical Demand: This dimension refers to the amount of physical effort required to accomplish a task.
- Temporal Demand: This dimension assesses how time-pressured the task is perceived to be.
- Effort: This dimension relates to the amount of mental and physical effort felt by the user during the performance of a task.
- Performance: This dimension assesses how the user evaluates their own performance on the task and whether they feel satisfied with their output or think they could have done better.
- Frustration: This dimension refers to the level of frustration or annoyance felt by the user while performing a task.

These dimensions allow for the assessment of different aspects of mental workload and to determine if the immersive experimental simulator induces a higher mental workload than the Cockpit simulator. This questionnaire has been used numerous times in the field of aeronautics and piloting (Shively et al. (1987), Battiste et al. (1988), Corwin et al. (1989), Nataupsky & Abbott (1987), Tsang & Johnson, (1989)). It is therefore highly suitable for our use case.

In addition to this questionnaire, a second questionnaire, the Workload Profile, was used.

- The "Partial" Workload Profile (Tsang & Velasquez 1996) for mental workload

The Workload Profile (WP) questionnaire, developed by Tsang and Velazquez (1996), is a tool for subjectively measuring mental workload criteria. It is based on Wickens' model of

1
2
3
4 multidimensional load (1980, 2002, 2004). Unlike the NASA TLX, its purpose is to quantify
5 the use of attentional resources utilized for the task. The questionnaire comprises a total of 7
6 dimensions, but for our purposes, only 4 have been retained as they correspond to the main
7 demands made on the flight task :

- 8 - Spatial Processing.
- 9 - Verbal Processing.
- 10 - Visual Processing.
- 11 - Manual Response.

12
13 These aspects of workload are complementary to the dimensions measured via the NASA TLX.

14 - *Simulator sickness measure* :

15 Simulator Sickness Questionnaire (SSQ; Kennedy 1993)

16 The SSQ was developed by Kennedy et al. (1993) and translated by Bouchard et al. (2009). In
17 this study, the three dimensions of the SSQ are measured through 16 questions on an adapted
18 scale from 0 to 100 : nausea, oculomotor discomfort, and disorientation. This questionnaire is
19 essential given the increased risk of simulator sickness with the use of virtual reality headsets.

20
21 - *Regarding the flight parameters retained to measure performance* :

22 Five physical data points are important to assess flight quality at the beginning of training :
23 time (in seconds), distance from the optimal trajectory (in meters), and representations of the
24 three axes of the aircraft: roll (in degrees), altitude (in meters), and heading (in degrees). To
25 handle the significant amount of data concerning the performance part, an automatic analysis
26 tool had to be developed. For each flight parameter, a performance was assigned in terms of
27 accuracy. For altitude, heading, and roll, an additional stability criterion (quantification of the
28 variation and the rate of change of the value) was considered. Finally, these data are compared
29 to those obtained by the reference professional pilot. They are thus represented as a percentage
30 of equivalence with respect to the pilot performance.

31 To define these criteria, this work referred to the instructor's guide published by ENAC (Ecole
32 Nationale de L'Aviation Civile ; AMDT N°00 - 17 NOV 2014, second Edition) which is the
33 French reference for flight certification. In addition to this, it relied on the recommendations of
34 the instructor consulted for the study, as well as on the observations and interviews carried out
35 during the observation phase of the first stages of flight by Rendy (2023), in his thesis work
36 about the analysis of piloting activity and the impact of immersion technology in a virtual
37 environment.
38
39

40 41 Results

42 *Data analysis*

43 To test our hypotheses, the results of each group of participants (corresponding to the two
44 simulators conditions) were compared. Statistical analyses were performed using the Jasp
45 Software (<https://jasp-stats.org/>). The statistic use for the decision was a two-factors analysis of
46 variance (ANOVA) with two independent groups and two repeated measures : $S22 < G2 > *$
47 Session2. Given that this analysis is not very sensitive to moderate deviations from normality
48 and considering that most of the distributions are normal or close to normal, this type of
49 analysis was favored. Indeed, studies using non-normal distributions have shown that the rate
50 of false positives is minimally affected by this violation of the assumption (Glass et al. 1972,
51 Harwell et al. 1992, Lix et al. 1996).

52 To address our hypotheses, three effects were analyzed, corresponding to the principal effects
53 and the interaction between our two factors, the type of simulator (2 groups of participants),
54 and the repetition of the flight training session (2 sessions).
55
56
57
58
59
60

Impact on Mental Workload

The analysis of the dimensions of the Nasa-TLX (table 4) and Workload Profile (table. 5) indicated no significant effect of the type of simulator and no significant interaction with the repetition of the sessions. Thus, the two simulators do not induce statistically different levels of mental workload throughout the learning process with close values for means and Standard deviations (Global Nasa-TLX Score for the Cockpit $M=44,16$ and $SD=14,59$ and for the Headset $M=48,86$ and $SD=6,68$). It is nevertheless interesting to note that, despite not statistically significant, the Workload Profile score of manual effort tends to be higher for the Headset simulator during the first session (fig.6).

/ Insert table 4 about here/

/ Insert table 5 about here/

In addition, for both simulators, training influenced the mental load experienced. More precisely, there was a reduction in the effort felt, as measured by Nasa-TLX (table 4 and fig. 4) and a reduction in the global score of WP (table 5, fig. 5) which was mostly explained by a decrease of manual and verbal demands felt between session 1 and 2 (table 5 and fig. 6).

/ Insert figure 4 about here/

/ Insert figure 5 about here/

/ Insert figure 6 about here/

Impact on Presence

Concerning the comparison between the two simulators, the SAM questionnaire reveals that the Cockpit simulator induced significantly higher presence (global score) than the Headset simulator (table 6). Essentially through the dimensions "Mental model of the virtual space", "Self-location" and "Possible actions" (fig. 7). This was completed by the QEP showing better haptic feeling for the Cockpit simulator (table 7 and fig. 8).

/ Insert table 6 about here/

/ Insert table 7 about here/

/ Insert figure 7 about here/

/ Insert figure 8 about here/

Concerning the evolution of the feeling of presence, the QEP reveals that it increased significantly from one session to another, whatever the simulator, mainly through the item's possibility to act, self-evaluation of performance and sounds (Table 14 and fig. 9).

/ Insert figure 9 about here/

Finally, it is important to note that the two simulators induce a fairly high level of presence (tables 6 and 7).

Impact on Simulator Sickness

There was no significant effect of the simulator or session repetition, nor any interaction effect (table 8). This lack of effect should be considered in light of the short time spent in the simulator (two 15-minute sessions with a maximum break in between).

/ Insert table 8 about here/

1
2
3
4 / Insert table 9 about here/
5

6 *Impact on Flight Performances*

7 Overall, the global performance score was not significantly different between the two
8 simulators (cf. Table 10).
9

10 / Insert table 10 about here/
11

12 However, regarding flight learning, there were aspects of performance that were more
13 important to consider than others. Indeed, the stability of the axes of rotation is a much more
14 critical safety point than time management. Looking at the overall index thus provides a general
15 idea of performance but is insufficient to properly qualify the quality of learning.

16 It was therefore more interesting to individually consider each of the measured performance
17 criteria to determine whether the learning differed based on these criteria.

18 Regarding time, the precision of all parameters, and heading stability, the two simulators could
19 not be considered different. The only flight variables for which the two simulators differed
20 were roll stability and altitude stability, with higher performance in the Cockpit simulator (fig.
21 10).
22
23

24 / Insert figure 10 about here/
25

26 Moreover, both simulators enabled significant improvement in performance with repeated
27 trials for almost all flight variables (time, distance accuracy, roll accuracy, altitude accuracy,
28 heading accuracy, and altitude stability, fig. 11). The only exception was for roll stability that
29 did not change significantly (table 8 and fig. 10).
30

31 / Insert figure 11 about here/
32

33 To ensure that both simulators enable equivalent learning, interaction effects must also be
34 considered (Table 10). It was observed only an interaction effect on the time factor. Figure 12
35 shows that this effect was due to the improvement of the headset group, whose time caught up
36 to that of the Cockpit group (whose time did not evolve from one session to another) during
37 the second session. (Table 11).
38
39

40 / Insert figure 12 about here/
41

42 / Insert table 11 about here/
43
44

45 Discussion

46 The primary objective of this study was to demonstrate that a Headset simulator could be as
47 effective as a traditional Cockpit simulator for training novice pilots to ab initio virtual fly.
48 More specifically, the study aimed to compare the subjective experience of presence and
49 mental workload for both types of simulators, ensuring that the Headset provided at least the
50 same level of presence as the Cockpit simulator. Additionally, it aimed to assess the learning
51 performance by measuring the improvement in flight skills during training on both simulators,
52 confirming that the Headset simulator supports effective learning. Furthermore, the study
53 sought to ensure that the Headset simulator did not cause more simulator sickness than the
54 Cockpit simulator, validating its usability. Ultimately, the goal was to leverage the benefits of
55 both Headset and traditional simulators to create a compact, cost-effective, and immersive
56 training tool. Until now, this had never been considered. In their work, Paul et al (2025) already
57
58
59
60

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

deplored this. Although these authors were able to demonstrate the potential for virtual reality and desktop simulators regarding training procedural task in ab initio fly training, they observed a limited learning capacity of this devices because of the lack of haptic feedback. Nevertheless, they did not test a mixed solution between VR headset and cockpit. In our study, we showed that full visual immersion with a headset combined with haptic feedback identified as essential for the task in question, even with a non-optimal level of fidelity, can enable learning equivalent to that of a simulator with a much higher level of fidelity.

The remainder of this section details the different contributions of our study on the processes at work when exposed to VR learning with different levels of fidelity.

Mental workload

Regarding the overall subjective mental load score, user perceptions do not appear to differ between the two simulators, considering the absence of statistically significant differences between the Nasa-TLX scores and close values for the observed means and standard deviations. These results are encouraging and tend to support our hypothesis H1 that the simulators would not differ in terms of induced mental workload. Obviously, these results are to consider with caution as non-significant results do not confirm equivalence between the two conditions. Moreover, it is noted that the perceived manual effort for the Headset simulator group seems qualitatively higher during the first session (despite these differences were not statistically significant). This could be related to the difference in the quality of manual controls. This is worth noting, as it could potentially shed light on the results obtained on other variables and it will be discussed in the following sections. In any case, after a second session with the Headset simulator, the level of manual workload qualitatively approaches that of the Cockpit despite the reduced quality of the controls. Thus, despite the existing differences in the technology used, the relevance of flight controls in terms of perceived mental load can be validated since the only difference (qualitatively observed) between the two simulators disappears with training.

Moreover, by examining the Workload Profile, it was found that, for both simulators, the perceived workload decreased from one session to the next, particularly for the verbal and manual dimensions. This was also observed for the effort dimension of the NASA TLX. This decrease is likely related to better situational mastery as well as the beginning of automation of certain piloting actions. This reduction of mental workload with mastery has often been observed, as in the study of Gaetan et al (2015), where expert helicopter pilots showed few mental workload when compared to novice. Thanks to this reduction in mental workload, in our study, the participant can allocate more attentional resources to piloting the plane. This is to be considered in relation with the significant increase in performance between the two sessions for both simulators.

Presence

The study evaluated user presence in the two types of simulators, the Headset simulator and the Cockpit simulator, using the Pictorial Presence Self-Assessment Manikins (SAM) and the Presence State Questionnaire (QEP). First it is noteworthy that both simulators induce a high level of presence despite not exactly the same in nature. If the Global score of QEP did not show a significant difference between the two simulators, the SAMs' dimensions revealed differences, challenging hypothesis H2 supposing that the Headset simulator would achieve the same level of presence as the Cockpit simulator.

This is the case concerning, the dimension "Mental Model of the Virtual Space". This dimension assesses the user's ability to form a mental representation of the virtual environment and their position within it. The study shows a significant difference between the Cockpit

1
2
3
4 simulator and the Headset simulator, with a lower score for the Headset. This should be
5 considered in conjunction with the "Self-Location" dimension which also showed a significant
6 difference, with a less accurate perception of bodily location in the Headset simulator. This
7 result could be due to the lack of a complete visual representation of the body in VR. It is
8 known that sense of embodiment (SOE), ie. the sensations of being inside, having, and
9 controlling a body (Guy, Normand, Jeunet-Kelway & Moreau, 2023) is crucial for achieving
10 good spatial presence in a virtual environment (Kilteni, Groten & Slater, 2012; Sanchez-Vives
11 & Slater, 2005). Avatar appearance is a crucial factor affecting SOE (Guy et al., 2023). In the
12 case of the Headset simulator, the users see only their hands and not their entire body. As a
13 result, they cannot transpose a complete body schema into the virtual environment. Especially,
14 this can influence the perception of the size and relative position of objects in the virtual
15 environment since the user does not have the reference of their own body's size. Zhao, Wang
16 & Wang (2023) showed that limited self-representation affect spatial presence and spatial
17 awareness during virtual locomotion. Additionally, users sometimes lose visual tracking of
18 their hands. This happens during a very short period of time. But this can lead to very short
19 interruptions in presence and the need for users to recalibrate the space to position and orientate
20 themselves in the virtual environment, which can impact their feeling of self-location but also
21 their perception of space relative to themselves and thus their mental model of virtual space.
22 Added to this is the fact that the lack of high-quality haptic feedback may have affected the
23 perception of distance and space due to less consistent information between the visual and
24 proprioceptive systems. This could be confirmed by the only item of the QEP showing a
25 difference between the two simulators, the haptic one. Nevertheless, numerous studies have
26 failed to show the superiority of haptic feedback over passive feedback for the sensation of
27 presence (ex. Kim et al., 2023; Viciano-Abad, Lecuona & Poyade, 2010) making this
28 interpretation less relevant.

29
30 However, the haptic limitation may help explain another result. The study revealed a significant
31 difference in the "Possible Actions" dimension between the Cockpit simulator and the Headset
32 simulator, with the latter being rated lower. This difference could be attributed to the technical
33 characteristics of each system. The Cockpit simulator offers high realism with many physical
34 flight controls, enhancing the user's feeling that he can act in the virtual space. In contrast, the
35 Headset simplifies the interface and offers less possibility of control, affecting the perception
36 of possible actions. Moreover, the haptic feedback technology of the Cockpit simulator might
37 increase this difference, with more advanced actuators probably improving the understanding
38 of interactions and the consequences of actions. This explanation is also congruent with the
39 qualitative higher feeling of manual effort noted during the first flying session for the Headset
40 simulator.

41
42 Finally, it is not surprising that haptic dimension is perceived as weaker in the case of the VR
43 headset, given that only passive-haptic feedback was used for the action buttons. Combined
44 with the lower fidelity of the haptic feedback from the actuators, this can lead to the feeling of
45 having less realistic haptic feedback regarding the actions performed. It should be noted,
46 however, that despite these limitations, the headset's haptic score remains quite good (more
47 than 60 over 100).

48
49 Nevertheless, it is noteworthy that only three dimensions of the SAM differed between the
50 simulators. By contrast, dimensions related to "Attention allocation", "Suspension of
51 disbelief", and "High-level cognitive demands" did not show a significant difference between
52 the two types of simulation, suggesting that hypothesis H2 is partially valid. The combination
53 of these three dimensions is interesting to consider, showing that whatever the technology used,
54 participants ignore the presence of the media (the "suspension of disbelief" threshold is high
55 in any cases) as long as they are captivated by the simulated flight situation (same level of
56 "Attention allocation" and High level of "Higher cognitive involvement").
57
58
59
60

Moreover, regarding the QEP, there are no significant differences between the two simulators on the global score and only one dimension over the seven considered showed a significant difference. These results tending to support hypothesis H2, of equivalence between the simulators, with regard to the dimensions of presence measured by this questionnaire. QEP also revealed that task repetition increases the global score of presence whatever the simulator. Users become more accustomed to the virtual environment, which enhances their familiarity and competence, thereby improving their sense of presence. The improvement of a dimension in particular, the “possibility to act”, is interesting. This result should be considered in the light of the changes we observed in mental workload over the course of the sessions. These decreased for both simulators, particularly with regard to the notions of effort and manual load, suggesting that the constraints linked to the manipulation of the actuators actually decreased with training, and could explain an improvement in the feeling of being able to act, and therefore, impact the overall feeling of presence.

To sum up, despite the differences noted between the simulators, the overall results show that both are comparable on a majority of presence dimensions. This suggests that the Headset simulator can be a viable alternative for some simulation applications.

Simulator sickness

Despite the use of the virtual reality headset, which is known to amplify simulator sickness, there is no significant difference between the two simulators. Therefore, hypothesis H3 can be validated.

However, in the Headset group, the observed values tended to slightly qualitatively increase after the second session despite not significantly different. Nevertheless, these values remain relatively low (no more than the quarter of the total scale), indicating that, for short usage periods, the Headset simulator remains usable and does not cause significantly more sickness than the Cockpit simulator. This is remarkable considering the known risks associated with this type of device and considering that participants were exposed for the first time to this technology. These results may seem high compared to the results observed in highly trained pilots using flight simulators. Indeed, the precept study by Kennedy et al (1993) on the SSQ indicated an average overall score of 9.8. To make it comparable to the scores observed in our study, we have to consider this score with respect to the maximum score that we can obtain on this scale, i.e. 4.2% of the maximum score (the highest overall values representing 8% of the maximum score). However, these values were obtained on fairly old simulators that were less realistic and immersive than those tested in our study. In addition, the pilots involved were highly trained and accustomed to using a flight simulator, and as the study showed, simulator sickness tends to diminish with pilot training. This is generally true for many virtual reality situations (Duzmanska et al, 2018). In a recent study, Nakdimon et al (2025) obtained a mean overall SSQ score of 32.7 (i.e. 14% of the maximum score) on helicopter simulators (CH-53 and UH-60 models) with trained Israeli Air Force pilots whereas, with participants not trained in flight simulators, Kim et al (2023) obtained much higher values in a fixed-base simulation condition relatively comparable to ours, with a mean overall score of 73.8 (i.e. 31.4% of the maximum score). In fact, comparison with other studies in the field remains difficult given the wide variety of protocols considered, the latter manipulating numerous factors likely to influence simulator performance (Martin Calderon et al, 2024; Duzmanska et al, 2018, Somerville et al, 2024): participants' expertise regarding piloting and flight simulators, duration of the test, nature of the flight exercise (more or less straight trajectory, changes in altitude, etc., Calderon et al, 2024). In the case of our study, the lap may be more likely to generate simulator sickness by inducing more changes in direction and altitude over a short period of time, especially since the participants were new to both aircraft piloting and flight simulators. In this context, the scores obtained on the headset simulator remain relatively low.

Performance and learning

The study compared user performance in the two simulators, focusing on different variables characterizing aircraft behavior. The results showed an improvement in performance over the sessions, indicating effective learning on both simulators with respect to the considered virtual flying task.

Regarding overall performance, the two simulators did not show a significant difference between the groups, tending to support hypothesis H4. However, when examining the dimensions related to aircraft stability, significantly less stable performance was observed in the group using the virtual reality headset. The difference in stability between the two simulators is attributed to the precision of the controllers, which are less accurate in the Headset simulator. Additionally, the haptic feedback provided by the Headset simulator does not achieve the same stability, as it can lock the yoke, making spring-based stability more difficult to obtain. These differences in actuators directly influence user performance. This is the criticism often opposed to many flight simulators (Marques et al, 2022). Nevertheless, the magnitude of the aircraft stability variability remains acceptable for initial learning. However, it can be assumed that this may become a problem later on and will need to be resolved. Fortunately, this is an element that can be easily modified subsequently to align more closely with the Cockpit simulator.

Despite the difference in stability, both simulators exhibit similar overall accuracy, and the results show significant and comparable improvements in most evaluated areas. Thus, we can conclude that, ultimately, the simulators do not differ significantly in terms of performance.

Conclusion

Our primary objective was to demonstrate that our simplified Headset simulator is suitable for the intended learning situation, being comparable to existing, more complex solutions, already used for training. In the field of ergonomics, our work shows that a simulator with limited fidelity, but equipped with the sensory information essential to the activity (this being identified on the basis of a work reporting an analysis of this activity), can achieve promising results in terms of presence, mental workload and performance compared to a more complex simulator designed for all phases of learning. Thus, it is possible to use the device for specific usage, while considering its limitations, notably by adopting motorized feedback flight controls that would allow comparable performances and presence feeling.

As well as comparing the two simulators, our study is contributing to our understanding of the processes at work in virtual reality by relating the technical parameters to their consequences, both on mental workload and on the feeling of presence, and by specifying which dimensions of presence are affected. Our study also considers how these processes evolve during the learning process. We have thus been able to demonstrate the value of combining different tools for measuring mental load and presence in order to achieve a more detailed understanding of these processes, which is relatively rare in studies looking at the evaluation of industrial prototypes.

Regarding the limitations of this study, it is noted that further validation with a larger number of participants is necessary. Moreover, in the near future, it will be crucial to explore the long-term effects of using the VR simulator, particularly about the stability of learning performance and, of course, learning transferability to real flight situations.

Lastly, the absence of a real training school and professional trainers can be considered a weakness in the study, as it would have been preferable to professionally train participants on both simulators following a complete training program to benefit from professional guidance and ensure realistic training conditions.

Acknowledgements

The authors are grateful to Maureen Follérou for proof reading this paper. They are also grateful to the Agence National for the Research and Technology (ANRT) for supporting this research work.

Ethical approval

The present study has received a positive decision from the non-interventional research ethics committee of the University of Nantes (CERNI) with the reference number n°27012022-2.

Data availability statement

The data are available at the following link : <https://uncloud.univ-nantes.fr/index.php/s/bdcnjogCk8asjXo>.

Disclosure statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

- Asobo Studio. (2020). Microsoft Flight Simulator [video game]. Xbox Game Studios. <https://www.xbox.com/games/microsoft-flight-simulator>.
- Battiste, V., & Bortolussi, M. (1988). Transport Pilot Workload: A Comparison of Two Subjective Techniques. *Proceedings of the Human Factors Society Annual Meeting*, 32(2), 150-154. <https://doi.org/10.1177/154193128803200232>
- Bouchard, S., St-Jacques, J., Renaud, P., & Wiederhold, B. (2009). Side effects of immersions in virtual reality for people suffering from anxiety disorders. *Journal of CyberTherapy & Rehabilitation*, 2, 127-137.
- Martin Calderon, C., Bae, J., Cao, S., & Barnett-Cowan, M. (2024). Navigating Simulator Sickness: The Effect of Flight Maneuvers in Fixed-Base Flight Simulators. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 68(1), 1674-1679. <https://doi.org/10.1177/10711813241260681> (Original work published 2024).
- Cegarra, J., & Morgado, N. (2009). Étude des propriétés de la version francophone du NASA-TLX. In 2009 5ème Colloque de Psychologie Ergonomique (pp. 233-239). Nice, EPIQUE.
- Corwin, W., Sandry-Garza, D., Biferno, M., P. B., & Logan, A. (1989). Assessment of Crew Workload Measurement Methods, Techniques and Procedures. Volume 1. Process, Methods and Results. 237
- Duzmanska N, Strojny P and Strojny A (2018) Can Simulator Sickness Be Avoided? A Review on Temporal Aspects of Simulator Sickness. *Front. Psychol.* 9:2132. doi: 10.3389/fpsyg.2018.02132
- Gaetan, S., Dousset, E., Marqueste, T., Bringoux, L., Bourdin, C., Vercher, J.-L., & Besson, P. (2015). Cognitive Workload and Psychophysiological Parameters During Multitask Activity in Helicopter Pilots. *Aerospace Medicine and Human Performance*, 86(12). <https://trid.trb.org/view/1376917>

- Glass, G. V., Peckham, P. D., & Sanders, J. R. (1972). Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research*, 42(3), 237-288. <https://doi.org/10.3102/00346543042003237>
- Guy, M., Normand, J.M., Jeunet-Kelway, C. & Moreau, G. (2023). The sense of embodiment in Virtual Reality and its assessment methods. *Frontiers in Virtual Reality*, 4, pp.1141683.10.3389/frvir.2023.1141683. hal-04323074.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index) : Results of Empirical and Theoretical Research. In P. A. Hancock & N. Meshkati (Éds.), *Advances in Psychology* (Vol. 52, p. 139-183). North-Holland. [https://doi.org/10.1016/S0166-4115\(08\)62386-9](https://doi.org/10.1016/S0166-4115(08)62386-9)
- Harwell, M. R., Rubinstein, E. N., Hayes, W. S., & Olds, C. C. (1992). Summarizing Monte Carlo Results in Methodological Research : The One- and Two-Factor Fixed Effects ANOVA Cases. *Journal of Educational Statistics*, 17(4), 315. <https://doi.org/10.2307/1165127>
- Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire : An Enhanced Method for Quantifying Simulator Sickness. *The International Journal of Aviation Psychology*, 3(3), 203-220. https://doi.org/10.1207/s15327108ijap0303_3
- Kilteni, K., Groten, R., & Slater, M. (2012). The Sense of Embodiment in Virtual Reality. *Presence - Teleoperators Virtual Environ.*, 21(4), 373-387. http://dx.doi.org/10.1162/PRES_a_00124
- Kim, J, Lee, SM, Son, H, Park, T (2023). Effects of unlimited angular motion cue and cue discrepancy on simulator sickness, *Expert Systems with Applications, Volume 213, Part C*, 119196, ISSN 0957-4174, <https://doi.org/10.1016/j.eswa.2022.119196>.
- Kim, D., Kim, Y., & Jo, D. (2023). Exploring the Effect of Virtual Environments on Passive Haptic Perception. *Applied Sciences*, 13(1), 299. <https://doi.org/10.3390/app13010299>.
- Koglbauer, I., Riesel, M., & Braunstingl, R. (2016). Positive effects of combined aircraft and simulator training on the acquisition of visual flight skills. *Cognition, Brain, Behavior*, 20(4), 309-318
https://www.researchgate.net/publication/312173853_Positive_effects_of_combined_aircraft_and_simulator_training_on_the_acquisition_of_visual_flight_skills
- Laminar Research. (2017). X-Plane 11 [video game]. Laminar Research.
- Lix, L. M., Keselman, J. C., & Keselman, H. J. (1996). Consequences of Assumption Violations Revisited : A Quantitative Review of Alternatives to the One-Way Analysis of Variance F Test. *Review of Educational Research*, 66(4), 579-619. <https://doi.org/10.3102/00346543066004579>
- Longo, L., Wickens, C. D., Hancock, P. A., & Hancock, G. M. (2022). Human Mental Workload: A Survey and a Novel Inclusive Definition. *Frontiers in psychology*, 13, 883321.
- Macchiarella, N. D., Arban, P. K., & Doherty, S. M. (2006). Transfer of training from flight training devices to flight for ab-initio pilots. *International Journal of Applied Aviation Studies*, 6(2), 299.
- Mansikka H, Virtanen K, Harris D. (2019). Comparison of NASA-TLX scale, modified Cooper-Harper scale and mean inter-beat interval as measures of pilot mental workload during simulated flight tasks. *Ergonomics*, 62(2):246-254. <https://doi.org/10.1080/00140139.2018.1471159>
- Marques, E., Carim, G., Campbell, C., & Lohmann, G. (2023). Ab Initio Flight Training: A Systematic Literature Review. *The International Journal of Aerospace Psychology*, 33(2), 99–119. <https://doi.org/10.1080/24721840.2022.2162405>
- Nakdimon I, Gordon B, Assa A, Ben-Ari O. Simulator sickness among helicopter pilots: Incidence, threshold and severity using the SSQ. *Journal of Vestibular Research*. 2025;35(1):4-8. doi:[10.3233/VES-230111](https://doi.org/10.3233/VES-230111)

- 1
2
3
4 Nataupsky, M., & Abbott, T. S. (1987, janvier 1). Comparison of workload measures on
5 computer-generated primary flight displays. Human Factors Society, Annual Meeting, New
6 York, NY. <https://ntrs.nasa.gov/citations/19880048216>
- 7 Oberhauser, M., Dreyer, D., Braunstingl, R., & Koglbauer, I. (2018). What's Real About
8 Virtual Reality Flight Simulation? : Comparing the Fidelity of a Virtual Reality With a
9 Conventional Flight Simulation Environment. *Aviation Psychology and Applied Human*
10 *Factors*, 8(1), 22-34. <https://doi.org/10.1027/2192-0923/a000134>
- 11 Paul, N., Moncion, B., & Cao, S. (2025). An experimental comparison on the effectiveness of
12 various levels of simulator fidelity on ab initio pilot training. *Ergonomics*, 1–17.
13 <https://doi.org/10.1080/00140139.2024.2449110>
- 14 Rendy, J. (2023). Développement d'un simulateur de vol immersif via une analyse de l'activité
15 de pilotage et impact de la technologie d'immersion dans un environnement virtuel sur le
16 sentiment de présence et l'apprentissage [Unpublished PHD Thesis].
- 17 Robillard, G., Bouchard, S., Renaud, P., & Cournoyer, L. (2002). Validation canadienne-
18 française de deux mesures importantes en réalité virtuelle : l'immersive tendencies
19 questionnaire et le presence questionnaire. In 25e congrès de la Société Québécoise pour la
20 Recherche en Psychologie (SQRP).
- 21 Sanchez-Vives, M. V., & Slater, M. (2005). From presence to consciousness through virtual
22 reality. *Nature Reviews Neuroscience*, 6(4), 332-339. <https://doi.org/10.1038/nrn1651>
- 23 Shively, R., Bortolussi, M., Battiste, V., Hart, S., Pepitone, D. D., & Matsumoto, J. H. (1987).
24 Inflight evaluation of pilot workload measures for rotorcraft research.
25 [https://www.semanticscholar.org/paper/Inflight-evaluation-of-pilot-workload-measures-for-
26 Shively-Bortolussi/caabf0efb04b88ece8d97074a3292d82253a9cf4](https://www.semanticscholar.org/paper/Inflight-evaluation-of-pilot-workload-measures-for-Shively-Bortolussi/caabf0efb04b88ece8d97074a3292d82253a9cf4)
- 27 Somerville, A., Joiner, K., Wild, G. (2025). Effects of Flight Experience or Simulator Exposure
28 on Simulator Sickness in Virtual Reality Flight Simulation. *Multimodal Technol. Interact.* 9,
29 24. <https://doi.org/10.3390/mti9030024>
- 30 Tsang, P.S., & Johnson, W.W. (1989). Cognitive demand in automation. *Aviation, Space, and*
31 *Environmental Medicine*, 60, 130–135.
- 32 Tsang, P.S., & Velazquez, V.L. (1996). Diagnosticity and multidimensional subjective
33 workload ratings. *Ergonomics* 39(3), 358–381.
- 34 Valdehita, Susana & Ramiro, Eva & García, Jesús & Puente, José. (2004). Evaluation of
35 subjective mental workload: a comparison of SWAT, NASA-TLX, and Workload Profile
36 Methods. *Applied Psychology*. 53. 61 - 86. 10.1111/j.1464-0597.2004.00161.x.
- 37 Viciano-Abad, R., Lecuona, A.R., & Poyade, M., (2010). The influence of passive haptic
38 feedback and difference interaction metaphors on presence and task performance. *Presence* 19
39 (3), 197–212.
- 40 Weibel, D., Schmutz, J., Pahud, O., & Wissmath, B. (2015). Measuring Spatial Presence :
41 Introducing and Validating the Pictorial Presence SAM. *Presence: Teleoperators and Virtual*
42 *Environments*, 24, 44-61. https://doi.org/10.1162/PRES_a_00214
- 43 Wickens, C. D. (1980). The Structure of Attentional Resources.
44 [https://www.semanticscholar.org/paper/The-Structure-of-Attentional-
45 ResourcesWickens/3db607fec63f0fb874e55b8230d6106fbc165a2e](https://www.semanticscholar.org/paper/The-Structure-of-Attentional-ResourcesWickens/3db607fec63f0fb874e55b8230d6106fbc165a2e)
- 46 Wickens, C. D. (2002). Multiple resources and performance prediction. *Theoretical Issues in*
47 *Ergonomic Science*, 3, 159-177. <https://doi.org/10.1080/14639220210123806>
- 48 Wickens, C. D. (2004). Multiple Resource Time Sharing Models. In *Handbook of Human*
49 *Factors and Ergonomics Methods*. CRC Press.
- 50 Wirth, W., Hartmann, T., Böcking, S., Vorderer, P., Klimmt, C. et al. (2007). A Process Model of the
51 Formation of Spatial Presence Experiences, *Media Psychology*, 9:3, 493-
52 525, DOI: [10.1080/15213260701283079](https://doi.org/10.1080/15213260701283079)
- 53
54
55
56
57
58
59
60

Wright, S., O'Hare, D. (2015). Can a glass cockpit display help (or hinder) performance of novices in simulated flight training? *Appl. Ergon.* 47, 292–299. <https://doi.org/10.1016/j.apergo.2014.10.017>

Witmer, B.G. & Singer, M.J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments.* 7 (3): 225–240. doi: <https://doi.org/10.1162/105474698565686>

Zhao, J., Wang, Z., & Wang, Y. (2023, October). Effects of Different Levels of Self-Representation on Spatial Awareness, Self-Presence and Spatial Presence during Virtual Locomotion. In *2023 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 4128-4133). IEEE.

Zheng Y, Lu Y, Jie Y, Fu S. Predicting Workload Experienced in a Flight Test by Measuring Workload in a Flight Simulator. *Aerosp Med Hum Perform.* 2019 Jul 1;90(7):618-623. doi: 10.3357/AMHP.5350.2019. PMID: 31227035.

Figures Captions and Alt Text

Figure 1 Caption: a) Top: Alsim Cockpit Simulator AL250. Bottom: visual environment b) Top: Experimental Headset Simulator. Bottom: visual environment.

Figure 1 Alt Text : The upper part of the figure represents the two simulators used in the study: the cockpit simulator and the headset simulator. The lower part shows the visual environments used for both simulators.

Figure 2 Caption : Circuit top view and visual landmarks (forest, tennis court, runway).

Figure 2 Alt Text : The figure shows a top view of the circuit with visual cues: the forest in the upper left part of the image, the runway in the middle right part and the tennis court in the lower right part of the image.

Figure 3 Caption: Top-down view of the downwind situation.

Figure 3 Alt Text : The figure shows a schematic top down view of the downwind situation with the respective positions of the plane, the runway and the tennis court.

Figure 4 Caption: Evolution of the Nasa TLX effort score with standard error bars.

Figure 4 Alt Text: The figure shows bars graphs (with standard error bars) representing the Nasa-TLX mean effort score for the session 1 Cockpit and headset conditions (on the left part of the graph) and the session 2 Cockpit and Headset conditions (on the right part of the graph). There was a significant reduction in the effort score from session 1 to 2.

Figure 5 Caption: Evolution of the Global Workload Profile (WP) score with standard error bars.

Figure 5 Alt Text : The figure shows bars graphs (with standard error bars) representing the Mean global WP score for the session 1 Cockpit and headset conditions (on the left part of the graph) and the session 2 Cockpit and Headset conditions (on the right part of the graph). There was a significant reduction of the WP global score from session 1 to 2 whatever the Simulator.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Figure 6 Caption: Evolution of the verbal and manual dimensions of the Workload Profile (WP) dimension with standard error bars.

Figure 6 Alt Text : The figure shows bars graphs (with standard error bars) representing the Mean Verbal and Manual WP score for the session 1 Cockpit and headset conditions (on the left part of the graph) and the session 2 Cockpit and Headset conditions (on the right part of the graph). There was a significant reduction of the Mean Verbal and Manual WP score from session 1 to 2.

Figure 7 Caption: SAM questionnaire. Statistically significant differences between the Cockpit and Headset simulator for Global score and the dimensions Spatial Situation Model, Self-Localization, and Possible Actions (with standard error bars).

Figure 7 Alt Text : The figure shows bars graphs (with standard error bars) representing the SAM questionnaire results for the mean Global score and the mean scores of the dimensions Spatial Situation Model, Self-Localization, and Possible Actions of the session 1 Cockpit and headset conditions and the session 2 Cockpit and Headset conditions. The mean Global score and the mean scores of the dimensions Spatial Situation Model, Self-Localization, and Possible Actions are significantly higher for the Cockpit simulator.

Figure 8 Caption: QEP questionnaire. Statistically significant differences between the Cockpit and Headset simulator for Haptic score (with standard error bars).

Figure 8 Alt Text : The figure shows bars graphs (with standard error bars) representing the mean QEP score for haptic dimension for the Cockpit condition session 1 and 2 (on the left part of the graph) and the Headset condition session 1 and 2 (on the right part of the graph). The mean Cockpit score is significantly higher than the one of the Headset whatever the session.

Figure 9 Caption: QEP questionnaire. Statistically significant differences between the Cockpit and Headset simulator for Global score and the dimensions Possibility to act, Self evaluation of Performance and Sounds (with standard error bars).

Figure 9 Alt Text : The figure shows bars graphs (with standard error bars) representing the QEP questionnaire results for the mean Global score and the mean scores of the dimensions Possibility to act, Self-Evaluation of Performance and Sounds for the session 1 Cockpit and headset conditions and the session 2 Cockpit and Headset conditions. the mean Global score and the mean scores of the dimensions Possibility to act, Self-Evaluation of Performance and Sounds are significantly higher for the session 2 than for the session 1 whatever the simulator.

Figure 10 Caption: Percentage of performance equivalence with respect to the expert, for the two simulators, concerning Roll and Altitude Stability (with standard error bars).

Figure 10 Alt Text : The figure shows bars graphs (with standard error bars) representing the percentage of performance equivalence for the two simulators in session 1 and in session 2 concerning Roll and Altitude Stability (respectively on the left side and right side of the graph). the scores of the Cockpit simulator are significantly higher for the two variables.

Figure 11 Caption: Percentage of performance equivalence with respect to the expert, for the two simulators and for dimensions which significantly improved between the two sessions (with standard error bars).

Figure 11 Alt Text : The figure shows bars graphs (with standard error bars) representing the percentage of performance equivalence for the two simulators in session 1 and in session 2 concerning Distance Precision, Roll Precision, Altitude Precision, Altitude Stability and Heading Precision. The scores of the Cockpit simulator are significantly higher for the two variables. The percentage of performance equivalence increase significantly from session 1 to 2 whatever the simulator for each dimension.

Figure 12 Caption: Evolution of Time Performance for the Cockpit and Headset simulator during the two sessions. This corresponds to the time taken to complete the task compared with the time taken by the expert. Being 80% of the time means that participant is either 20% slower OR faster than the expert.

Figure 12 Alt Text : The figure shows bars graphs (with standard error bars) representing the Percentage of Time Performance equivalence for the Cockpit simulator during session 1 and 2 (on the left part of the graph) and for the headset simulator during session 1 and 2 (on the right part of the graph). There was a significant improvement of time performance for the Headset simulator between session 1 and 2.

Tables Captions

Table 1: Technical Specifications of the Two Simulators

Table 2: Technical specifications of the force feedback of the AL250 a) and the Headset simulator b)

Table 3 : Summary of the experimental procedure.

Table 4 : ANOVA effects of Nasa-TLX dimensions. Effect size (assessed with Omega Square) is added in bracket for significant effect.

Table 5 : ANOVA effects of Workload Profile dimensions. Effect size (assessed with Omega Square) is added in bracket for significant effect.

Table 6: ANOVA effects for SAM dimensions. For the simulator factor, means for Headset and Cockpit simulators are indicated as follow : H(value) and C(value). Effect size (assessed with Omega Square) is added in bracket for significant effect.

Table 7: ANOVA effects for QEP dimensions. For the simulator factor, means for Headset and Cockpit simulators are indicated as follow : H(value) and C(value). Effect size (assessed with Omega Square) is added in bracket for significant effect

Table 8 : ANOVA of the overall index of the SSQ

Table 9 : Means and Standard errors of the 3 dimensions of the SSQ in the two sessions, for each simulator. The scores were computed on a scale from 0 to 100. SSQ-n corresponds to nausea score, SSQ-om to oculomotor diseases, SSQ-d to disorientation and SSQ-G to global score.

1
2
3
4 Table 10: ANOVA effects of the performance dimensions. Effect size (assessed with Omega
5 Square) is added in bracket for significant effect.
6

7 Table 11: PostHoc effects of the Time performance with Bonferroni test. Significant effect size
8 effect determined with Cohen's d test.
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

Table 1

	Visual field size	Virtual size	Resolution	Sound effect	Frequency image rendering	Haptics (button)	Haptics (Flight control)	Flight model (aircraft behaviour)
Cockpit	120°x115°	232°x42°	1920x1080 per projector	Stereo speaker, Alsim sound effect	60 Hz	Real buttons (real haptic feedback)	Motorised force feedback flight control	Alsim single-engine propeller
Headset	110°x108°	360°x360°	1440x1600 per eye	Stereo headphones, Alsim sound effect	90 Hz	Virtual buttons (pseudo-haptic feedback)	Spring-return flight control	Alsim single-engine propeller

Table 2

a)

Column Position (%) Push(-) / Pull(+)	Column Force (N)	Column Position (%) Rotation	Column Force (N)
-80	-135	-80	-156
-40	-68	-40	-78
0	0	0	0
40	68	40	78
80	135	80	156

b)

Column Position (%) Push(-) / Pull(+)	Column Force (N)	Column Position (%) Rotation	Column Force (N)
-0.8	-11,76	-0.8	-4,90
-0.4	-7,84	-0.4	-2,94
0	0	0	0
0,4	7,84	0,4	2,94
0.8	16,67	0.8	4,90

Table 3

Simulator	Cockpit		Headset	
Participants	10		12	
Session no.	1	2	1	2
Training video	yes	no	yes	no
Flight training	15 minutes	1 to 2 circuits	15 minutes	1 to 2 circuits
Test Flight	2 circuits	4 circuits	2 circuits	4 circuits
Questionnaire	x5 at the end of the session	x5 at the end of the session	x5 at the end of the session	x5 at the end of the session

Table 4

Effects		Global	Mental	Physical	Temporal	Effort	Performance	Frustration
Simulator	F	1,390	1,005	0,041	1,383	1,916	0,424	0,031
df=1 residue =20	p	0,252	0,328	0,841	0,253	0,182	0,523	0,862
Repetition	F	2,803	4,293	0,064	3,169	5,698	0,356	1,666
df=1 residue =20	p	0,110	0,051	0,803	0,09	<0,05* (0,08)	0,557	0,212
Interaction	F	2,306	2,16	0,747	2,847	1,825	0,075	0,205
df=1 residue =20	p	0,145	0,157	0,398	0,107	0,192	0,787	0,656

Table 5

Effects		Global	Spatial	Verbal	Visual	Manual
Simulator	F	0.354	2,137	1,090	2,002	0,375
df =1 residue =20	p	0.559	0,159	0,309	0,172	0,547
Repetition	F	4.768	0,411	5,657	0,443	7,741
df =1 residue =20	p	<0,05* (0,03)	0,529	<0,05* (0,03)	0,513	<0,05* (0,06)
Interaction	F	0.004	2,374	0,040	0,319	3,458
df =1 residue =20	p	0.951	0,139	0,843	0,579	0,078

Table 6

		Global	Attention Allocation	Spatial Situation Model	Self-location	Possible Actions	Higher cognitive Involvement	Suspension of disbelief
Simulator	F	7.361	0,158	8,685	6,472	7,722	1,939	0,051
Means		H (3,69) C (4,13)	H (2,67) C (2,80)	H (3,42) C (4,35)	H (4,29) C (4,85)	H (3,46) C (4,25)	H (4,42) C (4,70)	H (3,92) C (3,85)
df =1 residue =20	p	<0,05* (0,13)	0,695	<0,05* (0,15)	<0,05* (0,11)	<0,05* (0,14)	0,179	0,823
Repetition	F	0.945	0,507	1,191	0,003	0,593	0,313	0,199
df =1 residue =20	p	0.343	0,484	0,667	0,956	0,45	0,585	0,661
Interaction	F	1.653	0,507	0,012	0,374	0,19	0,313	2,433
df =1 residue =20	p	0.213	0,484	0,914	0,548	0,668	0,582	0,135

Table 7

		Global	Attention Allocation	Spatial Situation Model	Self-location	Possible Actions	Higher cognitive Involvement	Suspension of disbelief
Simulator	F	7.361	0,158	8,685	6,472	7,722	1,939	0,051
Means		H (3,69) C (4,13)	H (2,67) C (2,80)	H (3,42) C (4,35)	H (4,29) C (4,85)	H (3,46) C (4,25)	H (4,42) C (4,70)	H (3,92) C (3,85)
df =1 residue =20	p	<0,05* (0,13)	0,695	<0,05* (0,15)	<0,05* (0,11)	<0,05* (0,14)	0,179	0,823
Repetition	F	0.945	0,507	1,191	0,003	0,593	0,313	0,199
df =1 residue =20	p	0.343	0,484	0,667	0,956	0,45	0,585	0,661
Interaction	F	1.653	0,507	0,012	0,374	0,19	0,313	2,433
df =1 residue =20	p	0.213	0,484	0,914	0,548	0,668	0,582	0,135

Table 8

		Global
Simulateur	F	1.763
df =1 résidus = 20	p	0.199
Répétition	F	0.021
df =1 résidus = 20	p	0.886
Interaction	F	3.767
df =1 résidus = 20	p	0.067

Table 9

		Session 1				Session 2			
		SSQ-n	SSQ-om	SSQ-d	SSQ-G	SSQ-n	SSQ-om	SSQ-d	SSQ-G
Headset	Means	12,53	23,29	17,82	17,9	19,82	29,28	16,25	21,8
	SD	12,79	19,64	19,43	17,3	18,68	21,46	16,13	18,8
Cockpit	Means	14,59	16,58	9,46	13,5	9,73	13,21	5,36	9,4
	SD	21,44	13,42	14,31	16,4	7,77	12,16	5,42	8,5

Table 10

		Global	Time	Distance Precision	Rolls precision	Rolls stability	Altitude precision	Altitude stability	Heading precision	Heading stability
Simulator	F	1.375	0,829	0,119	9,20E-04	23,989	0,786	7,382	1,502	1,21E-06
	df =1 résidus = 20 P	0.255	0,373	0,734	0,976	<0,001** (0,35)	0,386	<0,05* (0,13)	0,235	0,999
Repetition	F	44.300	12,792	8,819	22,946	0,492	8,806	9,543	25,338	4,172
	df =1 résidus = 20 P	< .001**	<0,05* (0,09)	<0,05* (0,11)	<0,001** (0,21)	0,491	<0,05* (0,10)	<0,05* (0,07)	<0,001** (0,15)	0,055
Interaction	F	1.032	5,714	0,001	0,178	0,825	0,002	0,839	0,223	1,154
	df =1 résidus = 20 P	0.322	<0,05* (0,04)	0,97	0,677	0,375	0,961	0,371	0,642	0,295

Table 11

Post Hoc Comparisons - Groupe * RM Factor 1

		Mean Difference	SE	t	Cohen's d	p _{bonf}
Cabine, Level 1	VR, Level 1	8.776	4.848	1.810	0.775	0.487
	Cabine, Level 2	-2.357	2.935	-0.803	-0.208	1.000
	VR, Level 2	-3.082	4.848	-0.636	-0.272	1.000
VR, Level 1	Cabine, Level 2	-11.133	4.848	-2.296	-0.983	0.177
	VR, Level 2	-11.857	2.679	-4.426	-1.047	0.002
Cabine, Level 2	VR, Level 2	-0.725	4.848	-0.149	-0.064	1.000

Note. P-value adjusted for comparing a family of 6

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Figure 1

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60



Figure 2

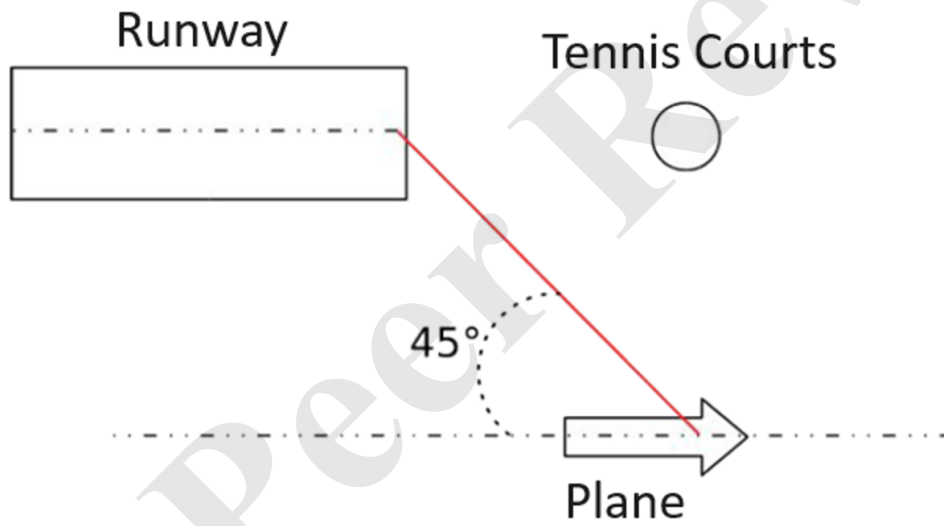


Figure 3

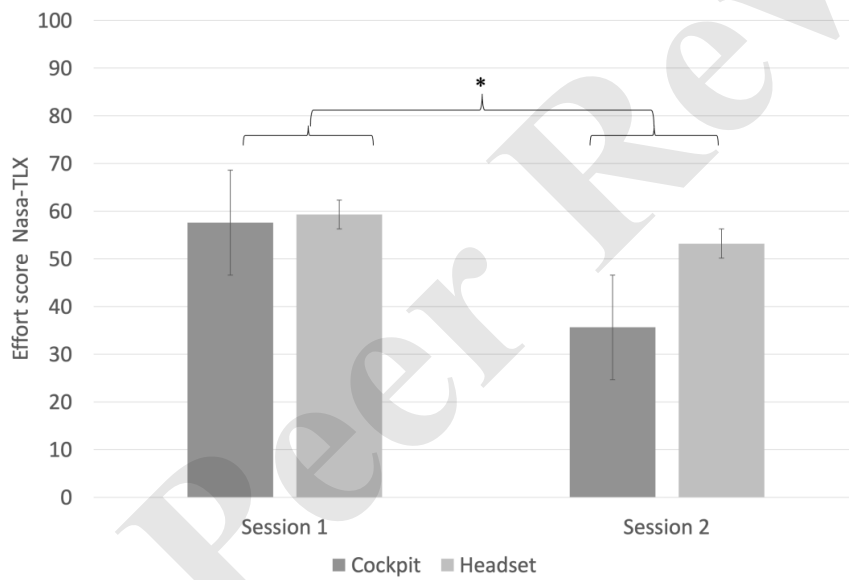


Figure 4

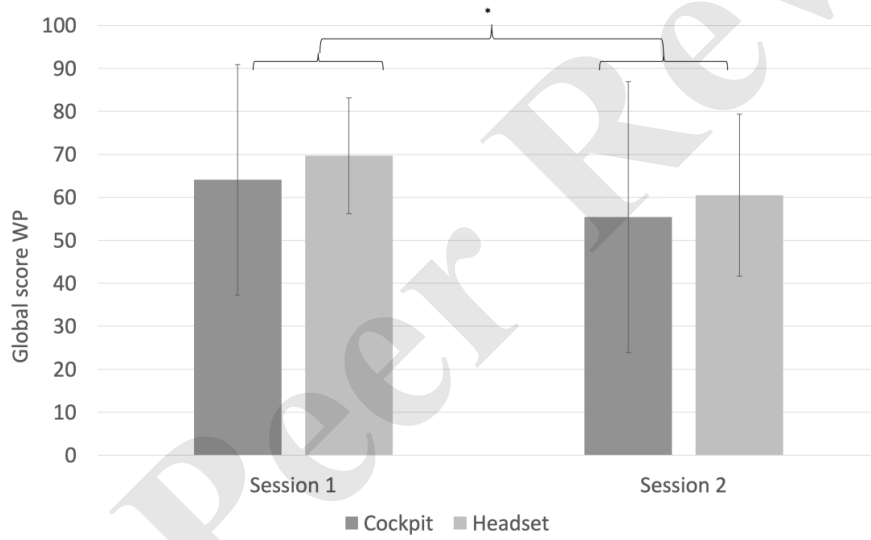


Figure 5

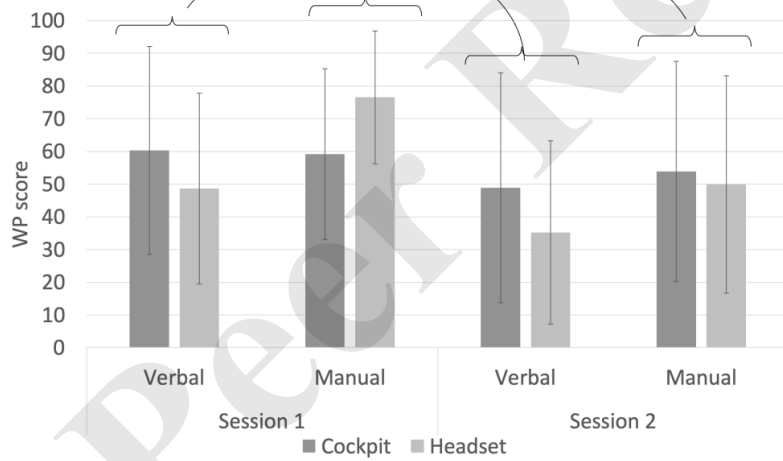


Figure 6

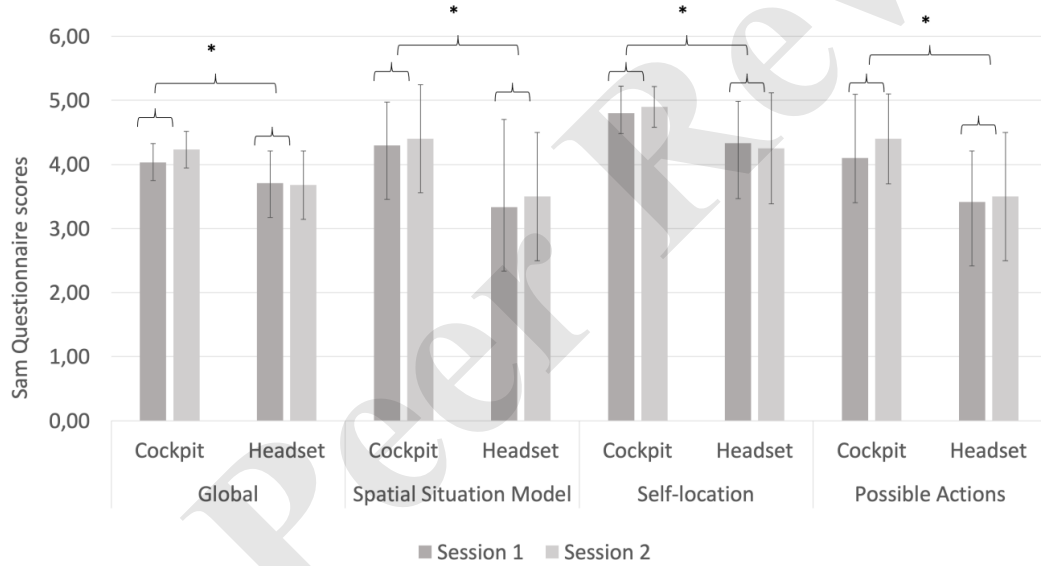


Figure 7

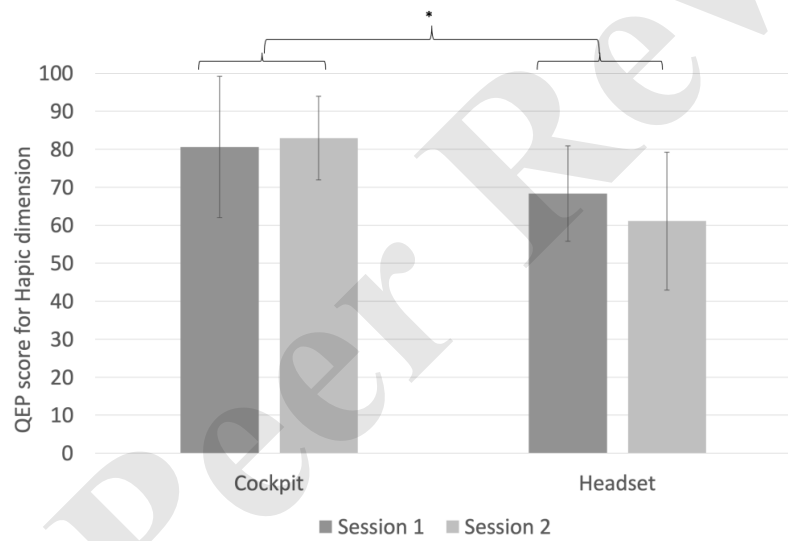


Figure 8

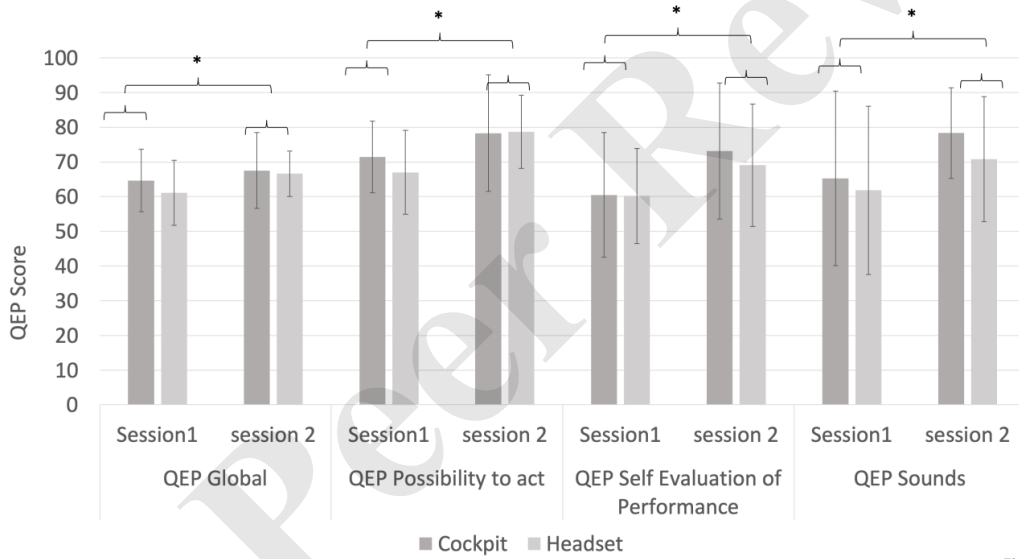


Figure 9

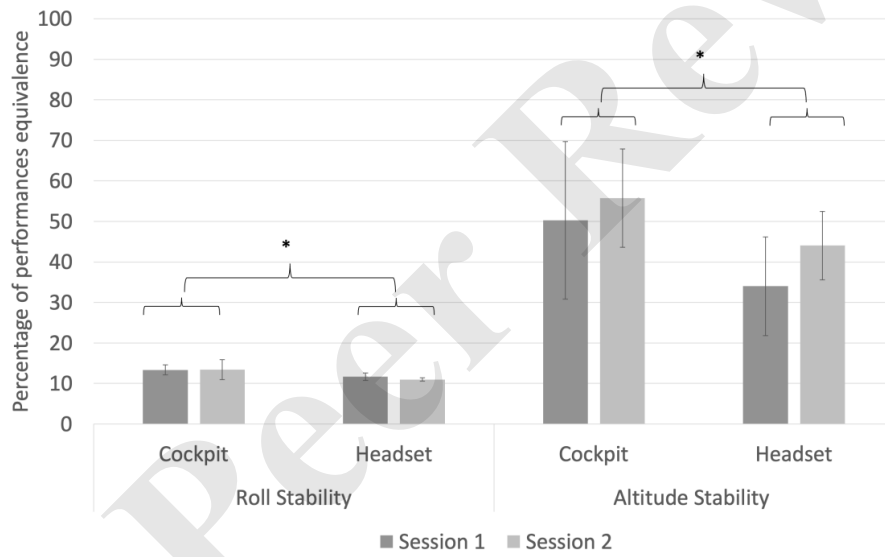


Figure 10

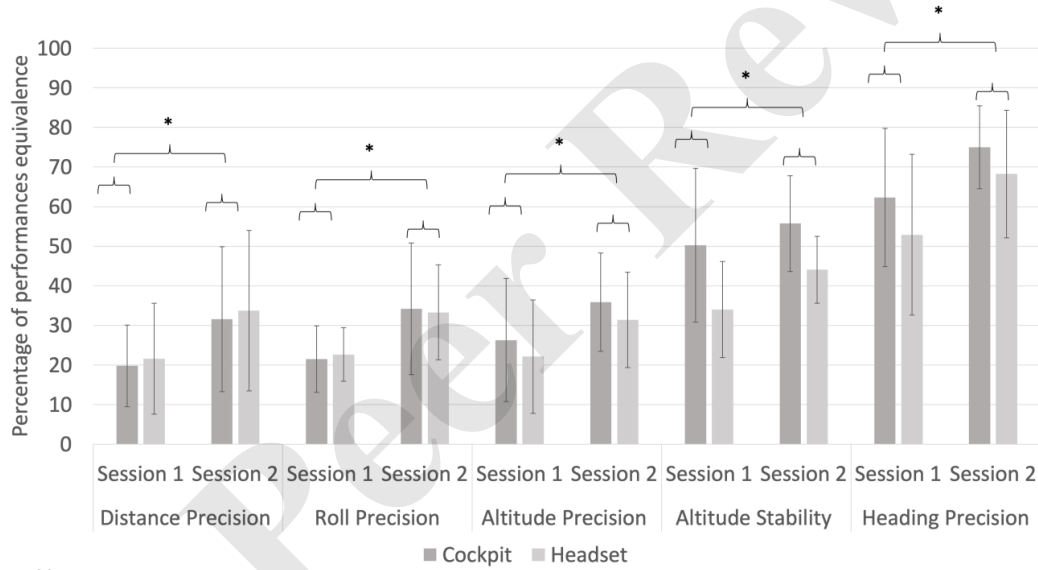


Figure 11

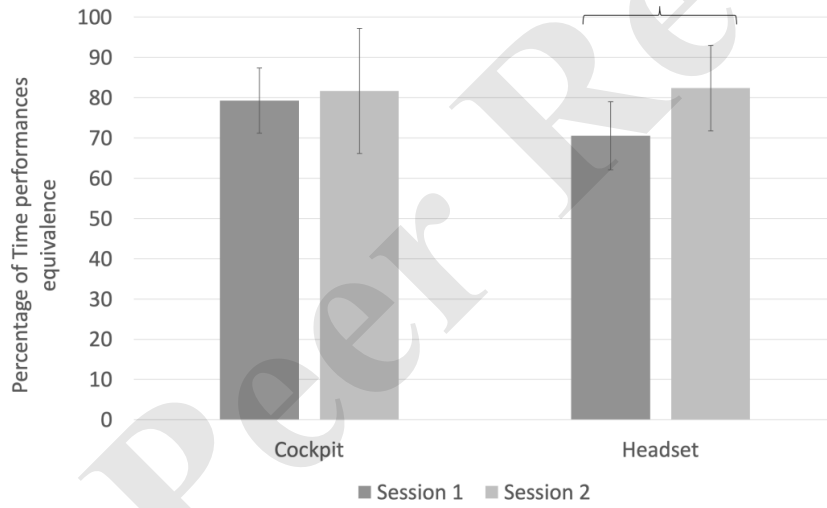


Figure 12