Comparison of Cardiac Activity and Subjective Measures During Virtual Reality and Real Aircraft Flight

Patrice Labedan, Frédéric Dehais, Vsevolod Peysakhovich

Comparison of Cardiac Activity and Subjective Measures During Virtual Reality and Real Aircraft Flight

Patrice Labedan(✉), Frédéric Dehais, and Vsevolod Peysakhovich

ISAE-SUPAERO, Université de Toulouse, Toulouse, France
patrice.labedan@isae.fr

Abstract. Pilot training requires significant resources, both material and human. Immersive virtual reality is a good way to reduce costs and get around the lack of resources availability. However, the effectiveness of virtual flight simulation has not yet been fully assessed, in particular, using physiological measures. In this study, 10 pilots performed standard traffic patterns on both real aircraft (DR400) and its virtual simulation (in head-mounted device and motion platform). We used subjective measures through questionnaires of immersion, presence, and ability to control the aircraft, and objective measures using heart rate, and heart rate variability. The results showed that the pilots were able to fully control the aircraft. Points to improve include updating the hardware (better display resolution and hand tracking) and the simulator dynamics for modelling ground effect. During the real experience, the overall heart rate (HR) was higher (+20 bpm on average), and the heart rate variability (HRV) was lower compared to the virtual experience. The flight phases in both virtual and real flights induced similar cardiac responses with more mental efforts during take-off and landing compared to the downwind phase. Overall, our findings indicate that virtual flight reproduces real flight and can be used for pilot training. However, replacing pilot training with exclusively virtual flight hours seems utopian at this point. AQ1 AQ2

Keywords: Virtual reality · Flight simulation · Heart rate · Heart rate variability · Piloting

1 Introduction

Despite a significant passenger traffic decrease in aviation due to the COVID-19 pandemic, the global air transport sector is recovering. The post-COVID-19 aviation industry will inevitably face challenges of pilot and instructors shortage. Pilot training requires significant resources, both material and human, such as flight simulators, flight instructors, and, of course, aircraft. It represents a high cost and generates constraints on the availability of these means. Safety is also a paramount aspect of pilot training, particularly, while flying on real aircraft (breakdowns, weather phenomena, etc.).

Immersive virtual reality (VR) seems to be an adequate alternative to reduce costs and get around the lack of availability of resources (aircraft, simulators, instructors) for skill acquisition [19,44]. The recent development makes the design of virtual environments such as cockpit or flight simulators more flexible [1]. The VR is already used

A. A. de Sousa et al. (Eds.): VISIGRAPP 2021, CCIS 1691, pp. 1–20, 2023.
https://doi.org/10.1007/978-3-031-25477-2_6

professionally or evaluated for training in various fields such as UAV piloting [36], fire fighting [6, 7], maritime domain [29], first responsers training [24], evacuation training [12], or mining industry training [45].

The medical field is a forerunner in the field of VR adoption [5, 20, 22, 28, 34]. VR now allows to deliver, in some cases, cost-effective, repeatable, standardized clinical training on demands. It is a powerful educational tool and implementation is growing worldwide [33]. In surgery, the use of virtual reality has been successfully tested for many years [38]. Surgeons in training can acquire skills without threatening the lives of patients, especially in laparoscopic surgery [16], with positive results in terms of feelings of presence and the ability to transfer the training to the real operation. This is also the case in cataract surgery [40], in oral and maxillofacial surgery [3], etc. These promising results in the field of surgery and its similarities to flight, including high levels of stress, accuracy, and risk-taking [35], make VR worthy of consideration for pilot training.

However, the use of VR as an operational learning tool still presents challenges [13] in terms of immersion, sense of presence [42], fatigue, and motion sickness [26]. Indeed, it is recognized that simulators do not reproduce the level of engagement that pilots may experience in real-world conditions [14, 31]. Studies comparing VR and simulator training [2, 27, 31] or simulator and real flight training [17], have already been conducted. A recent study [26] with pilot instructors showed that the strong feeling of immersion, combined with good controllability of the aircraft, generates high presence levels. Another study [32] showed that VR is an efficient tool for learning checklists in the early stages of pilot training. But recently it was shown that virtual reality flight simulations induce higher workload, physical demand, and effort, exceeding acceptable levels [2]. To date, to our best knowledge, no research except one has been found directly comparing VR and real flights. This study [25], with 4 participants, showed preliminary results with a higher heart rate and a lower heart rate variability in real flight compared with virtual reality. An interesting perspective for such a comparison is to measure subjective and objective indicators of the mental effort of pilots in both virtual and real flight situations. Cardiac activity, in particular, is a possible indicator for cognitive load [30], even in operational conditions [37]. A similar approach had already been carried out to compare simulators to virtual reality [27] and disclosed a slightly higher heart rate in virtual reality than in a flight simulator.

The present study focuses on the data acquisition and analysis of heart activity parameters, in real and virtual flight conditions, with student pilots in training. We also report subjective measures of pilots' experience in virtual reality, particularly the feeling of presence and the difficulty perceived by the pilots to perform different actions during the flight of the chosen scenario.

We build up upon a previous preliminary experiment with a limited participant number (N = 4) [25]. The present work includes 10 pilots, thus, allowing to perform a statistical analysis between different conditions and flight phases. Previous insights can be thus statistically verified. We also added other subjective questionnaires such as the Presence Questionnaire, the Immersive Tendencies Questionnaire, and the Flight Difficulty Questionnaire to evaluate the difficulties to perform the virtual reality flight (Fig. 1). AQ3

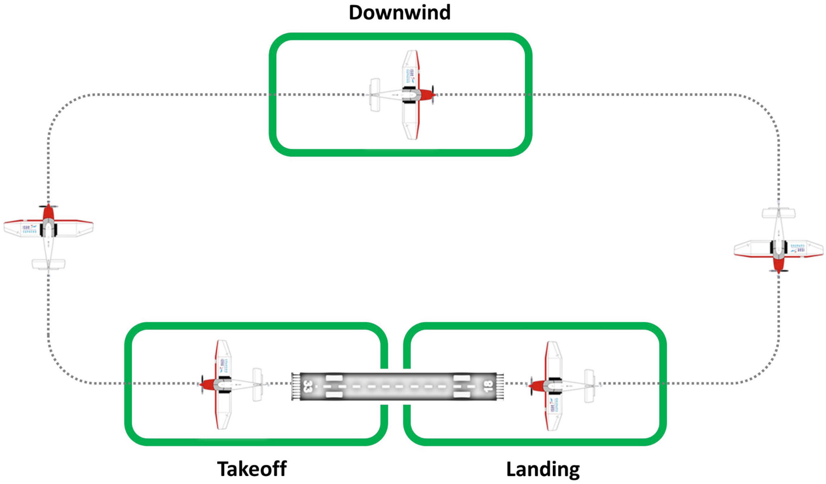

Fig. 1. The three flight phases: take-off, downwind and landing, each lasting 60 s.

2 Materials and Methods

2.1 Participants

Ten student pilots of the ISAE-SUPAERO, Toulouse, France, holding a private pilot's license (PPL) or in training to obtain it, participated in the experiment (all males, mean age 22.4 ± 4.4 years; mean flight experience 154 h). All were in good health, as evidenced by their flight medical certificates. No participant had a history of cardiac or neurological disease and, as required by aviation regulations, no participant was taking psychoactive substances or medications. Participants signed a consent form prior to the experiment. The virtual reality experience was approved by the Ethics and Research Committee of Toulouse (n2019-193). The flight experience was approved by the European Aviation Safety Agency with the permit to fly N°0147/21/NO/NAV (the flight conditions are defined in the document Monitoring Pilot's brain activity in real flights dated 26/01/2021, approved by EASA 60077217 on 24/03/2021).

2.2 Real Aircraft

The aircraft used during the experiments was the ISAE-SUPAERO experimental four-seater airplane Robin DR400 (160 HP). It is the same aircraft that is used to train our volunteers who participated in the study (Fig. 2). The experiment took place at Toulouse-Lasbordes airfield, in France (OACI code: LFCL).

2.3 Virtual Aircraft

We used the VRtigo platform (Fig. 4) [25,26], a VR flight simulator, that realistically reproduces the DR400 cockpit and Lasbordes Airfield environment to allow comparison with the real flight settings. This platform is composed of:

Fig. 2. The aircraft used for the experiment (Robin DR400).

Fig. 3. DR400 cockpit panel identically reproduced in virtual reality.

- Aerofly FS2 flight simulation software (IPACS);
- Runway 33 of Toulouse-Lasbordes airfield (OACI code: LFCL), identically reproduced, with some buildings commonly used for visual cues;
- DR400 panel, also identically reproduced to the aircraft used during real flights (Fig. 3);
- The 6-axis motion platform MotionSystems PS-6TM-150 with the following characteristics: heave –106.9 mm +117.1 mm, pitch –25° +25.6°, roll ±26°, yaw ±22.5°, surge –100 mm +121 mm, sway –99.5 mm +121mm;
- Simple and conventional controls: stick, rudder, throttle, and flap lever;
- A cockpit, including the controls and a pilot's seat;
- A virtual reality headset (HTC Vive);
- An Alienware "VR ready" Laptop computer.

Fig. 4. VRtigo: the virtual flight simulator at ISAE-SUPAERO.

2.4 Flight Scenario

The scenario consisted of three consecutive standard traffic patterns. This exercise is highly formalized [8,37] in terms of flight procedures and flight path which makes it a relevant candidate for comparing pilot's behavior in the two experimental conditions (i.e. VR vs real flight). The scenario was identical in both virtual and real flights, and consisted of three traffic patterns, with a touch-and-go between each of them. After the last pattern, the pilots made a final landing and complete stop of the aircraft. In this study, we focused our analyses on three specific phases of the standard traffic pattern:

– Take-off (from maximum power setting);
– Downwind (the center of the return path);
– Landing (before touchdown).

The separation into 60-s phases allowed to compare with the previous study performed in similar conditions [37], and to improve our previous study [25].

2.5 Measures

Cardiac Activity. The cardiac data were acquired with a Faros 360 eMotion electrocardiogram (ECG) device. It provides the raw ECG signal at a sampling rate 500 Hz and R-R intervals data using a built-in R-detection algorithm. To improve the signal quality, we applied a conductive gel to each of the three electrodes connected to the Faros system. Physiological measures were synchronized with the flight parameters with the Lab Streaming Layer (LSL) [23].

Flight Parameters. In the VR flights, the following Aerofly FS2 simulator flight parameters were streamed, recorded, and stored via the LabRecorder:

- Altitude (feet)
- Pitch and bank angle (degrees)
- Groundspeed (knots)
- Indicated airspeed (knots)
- Longitude/Lattiude (degrees)
- Mach number (feet)
- Magnetic and true heading (degrees)
- Throttle (%)
- Vertical airspeed (feet/min)

For the real flights, an ILevil 3-10-AW acquisition unit was used to collect the trajectory (via GPS), accelerations, altitude (in feet), speed (in knots), and yaw/pitch/roll information. Similarly, these data were recorded and stored via the LabRecorder. This acquisition unit had to be mounted in the aircraft's cargo area at a specific location that guaranteed the accuracy of the attitude data (roll, pitch, and yaw). These parameters were then used to automatically identify the three flight phases of interest.

Questionnaires. Regarding the ten pilots, four participated in the experiment without filling out the questionnaires (2019). The other six (2021) participated in the experiment with three subjective questionnaires about the virtual environment. All the questionnaires used a visual analog scale from 1 to 7 for each answer, with different significations according to the question.

The Presence Questionnaire (PQ). This questionnaire measures the feeling of presence in a virtual environment. There are many versions of it and the one we used is composed of 17 questions. It is inspired by the 2002 french version from The Cyberpsychology Lab of UQO (Université du Quebec en Outaouais) [4]. The signification of the answer on the visual scale from 1 to 7 is different for each question (please refer to the 2002 Cyberpsychology Lab version for more details);

The Immersive Tendencies Questionnaire (ITQ). This questionnaire measures differences in the tendencies of individuals to experience presence. We selected the 2002 french version from The Cyberpsychology Lab of UQO (Université du Quebec en Outaouais) [4]. This version is inspired by Witmer and Singer's 1998 original version

[43]. It is composed of 18 questions that measure the level to which the individual can cut off from external distractions to concentrate on different tasks. The participant must answer each question on a 7-point scale. A score from 1 to 7 is associated with each answer and the global score is calculated. These 18 questions are also divided into 4 subscales, which measure different aspects of immersion propensity (Focus, Involvement, Emotion, and Game). The "Focus" subscale measures the sustained attention generated by an activity (5 items). The "Emotion" subscale deals with the individual's ease of feeling intense emotions evoked by the activity (4 items). The "Game" subscale refers to the individual's ability to project him/herself into a playful context (video game, etc.) (3 items). The "Involvement" subscale measures the tendency of an individual to identify with characters or to feel completely absorbed by an activity (4 items). A score is calculated for each subscale. The Cyberpsychology Lab has established certain norms (minimum scores for the overall and per subscale). Like the presence questionnaire, the answer on the visual scale had different signification according to the question (more details in the 2002 Cyberpsychology Lab version [4]);

A Flight Difficulty Questionnaire (FDQ). A set of 14 questions, specific to aircraft piloting in our flight scenario, was used to assess and compare the level of difficulty of the different flight segments across conditions. This questionnaire was created in 2019 in our lab and was already used for a previous study [26]. For questions 1 to 4, the answer on the visual scale could be from 1 (not at all) to 7 (completely). For questions 5 to 14, the answer could be from 1 (not similar) to 7 (very similar). See Fig. 8 for details about the 14 questions.

2.6 Experimental Protocol

Real Flight. Three people were present on the aircraft: the left seated pilot-participant, the right seated flight instructor (FI) acting as a safety pilot (right seated), and the experimenter (backseater). Before getting on the plane, the participants received a briefing about the experiment and completed their pre-flight inspection. The experiment then placed the ECG electrodes on their torso.

During the flight, an LSL's Viewer application displayed ECG data and flight parameters in real-time, which was necessary for the experimenter to ensure data consistency over time. The first data check was performed between engine start and taxi. The experimenter started the ECG and flight parameters data recordings before the aircraft's first take-off and stopped after the last landing. The meteorological conditions were compatible with the flights. In most cases, the Ceiling and Visibility were OK (CAVOK), and the wind was calm. For some subjects, the wind was stronger, with 15 knots gusting 25 knots, with a 30 angle from the runway's axis. Temperatures sometimes rose to 36 °C in the cockpit on the ground.

Virtual Flight. The virtual reality flights were conducted in a temperature-controlled room. Three people were inside the room: the participant (on the VRtigo platform), the experimenter (monitoring ECG data and aircraft configuration), and the safety technician (controls the correct functioning of the moving platform, ready to interrupt the

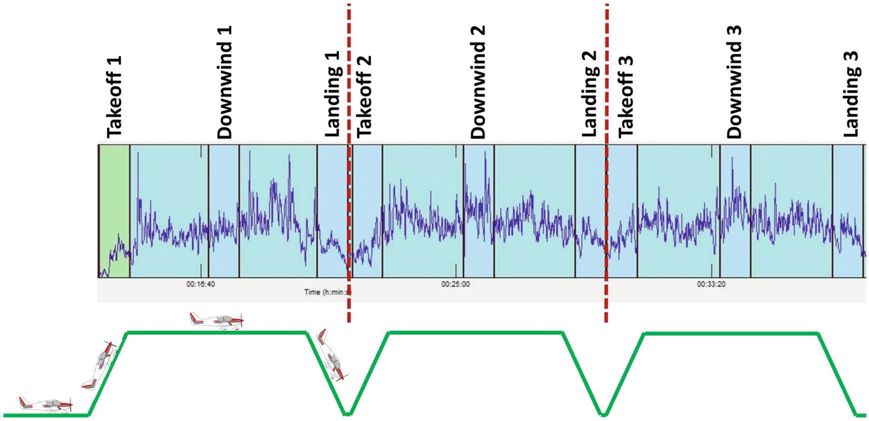

Fig. 5. The three flight phases on each of the three traffic patterns, over the R-R data for a complete flight (a 25-min. duration in this example). The green line schematically depicts the aircraft's altitude. (Color figure online)

simulation at any time by pressing an emergency stop button). The weather conditions were CAVOK, and no wind was programmed. The recordings were switched on and off at the same time as the real flights.

2.7 Data Analyses

Electrocardiogram. The R-R intervals of the raw ECG signal were detected using the built-in QRS detection algorithm of Kubios HRV software [39]. All the recordings were visually inspected to correct for potentially missed or false positive R-peak detection. We then computed the mean values of heart rate (HR; in beats per minute) and heart rate variability (HRV; assessed as Root Mean Square of the Successive Differences of the R-R intervals – RMSSD in ms, and NN50 and pNN50 – number and percentage, respectively, of R-R intervals that differ from each other by more than 50 ms) within the 60-s window of the three phases (take-off, downwind, landing) of the three traffic patterns (Fig. 5). Usually, studies report a decrease in HR and an increase in HRV (i.e. higher variability) as task demand gets lower [10,37,41].

The Kubios software also computes other metrics such as LF, HF, LF/HF in the frequency domain, but their computations on short-term signals are not recommended [18], therefore, given the 60-s window length, we only considered HR, RMSSD, NN50, and pNN50 metrics.

Statistical Analyses. The statistical analyses were carried out with JASP 0.16.1 software. Two-way (3 flight phases $\times$ 2 settings) repeated analyses of variances (ANOVAs) were computed over the HR and HRV metrics. The Greenhouse-Gessier sphericity correction was applied and the Holm correction was used for all post hoc comparisons. The significance level was set at $p < .05$ for all analyses.

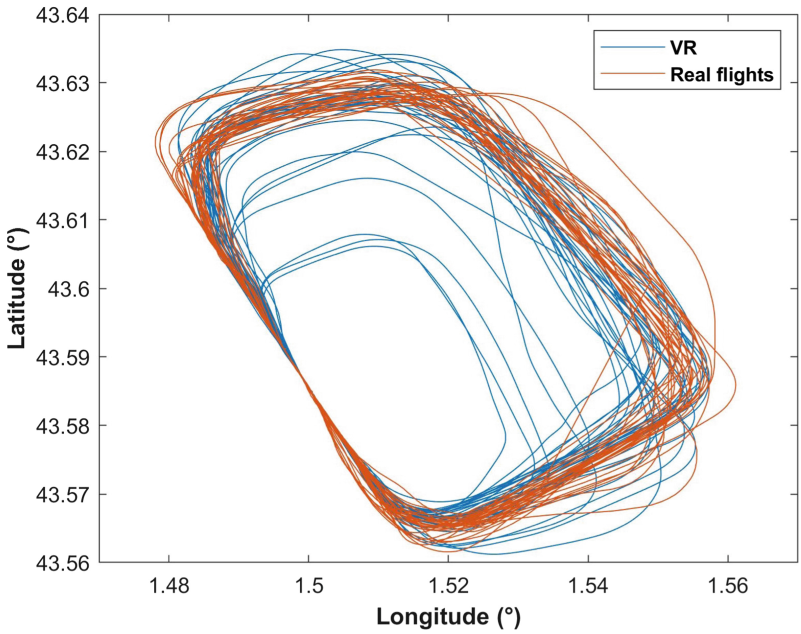

Fig. 6. Cumulative trajectories in virtual reality (blue) and real flights (orange). (Color figure online)

3 Results

3.1 Flight Parameters

The flight parameters have not been fully exploited for the moment. They were used in this study for the extraction of the three flight phases by the analysis of the following parameters: longitude, latitude, altitude, heading, and power. We then only visually checked the trajectories to verify their coherence between real and virtual reality flights (Fig. 6).

3.2 Subjective Measures

The Presence Questionnaire. Figure 7 represents the results of the Presence Questionnaire. The min score is 2.14 for question 10 ("How much of a delay did you feel between your actions and their consequences?"). The response could be from 1 ("no delay") to 7 ("long delay"). The max score is 6.57 for question 9 ("How involved were you in the experience in the virtual environment?"). The response could be from 1 ("no at all") to 7 ("completely").

The Immersive Tendencies Questionnaire. Participants showed a global mean score of 79.00 (SD = 11.98), and mean scores of 24.67 (SD = 4.37) for the Focus subscale, 21.42 (SD = 5.12) for the Involvement subscale, 15.75 (SD = 4.05) for the Emotion subscale, and 11.67 (SD = 2.94) for the Games subscale.

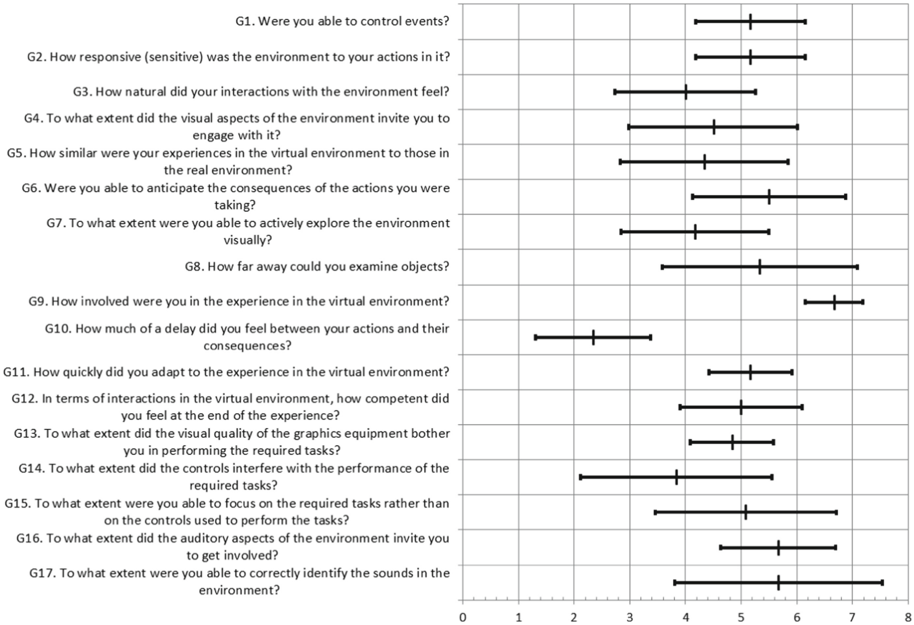

Fig. 7. Results of the presence questionnaire.

Means and standard deviations for the ITQ global score and subscales are presented in Table 1, along with norms from The Cyberpsychology Lab of UQO [4];

Table 1. ITQ results score (global and sub-items).

	Score	Norms
Global	**79.00 ± 11.98**	**64.11 ± 13.11**
Focus	24.67 ± 4.37	24.81 ± 7.54
Involvement	21.42 ± 5.12	15.33 ± 8.67
Emotion	15.75 ± 4.05	14.25 ± 6.70
Games	11.67 ± 2.94	6.56 ± 4.95

The Flight Difficulty Questionnaire. Figure 8 shows the results of the Flight Difficulty Questionnaire. The min score is 3.43 for question 14 (“How similar were your experiences during the landing phase in the virtual environment to those in the real environment?”). The response could be from 1 (“not similar”) to 7 (“very similar”). The max score is 6.57 for question 4 (“How well were you able to control the thrust of the aircraft?”). The response could be from 1 (“not at all”) to 7 (“completely”).

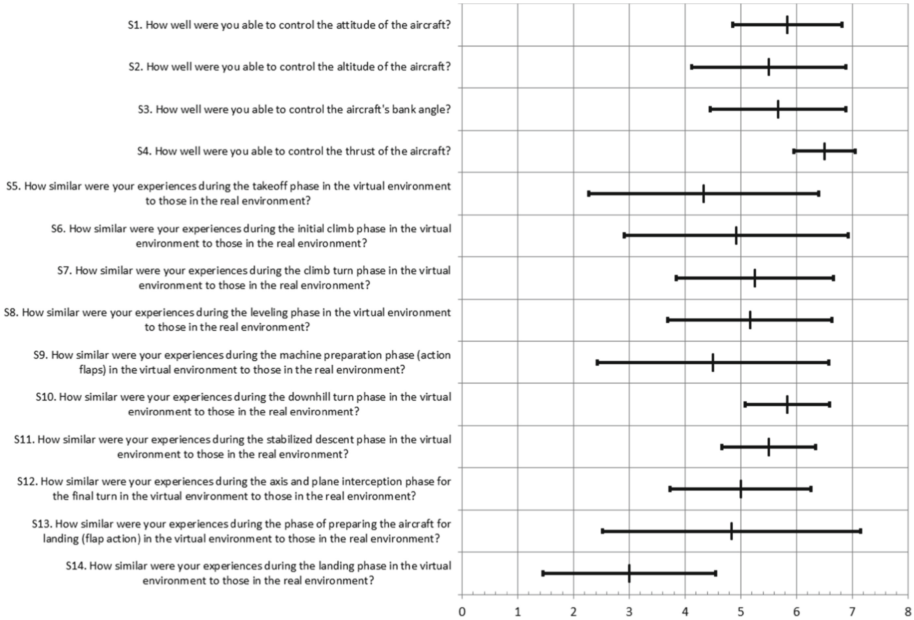

Fig. 8. Results of the flight difficulty questionnaire.

3.3 Physiological Measures

Heart Rate. A first two-way repeated ANOVA disclosed a main effect of the flight setting on HR, $F(1, 9) = 14.0, p = .005, \eta_p^2 = 0.610$, and a main effect of the flight phase on HR, $F(2, 18) = 19.1, p < .001, \eta_p^2 = 0.680$, as well as a significant flight setting × phase interaction, $F(2, 18) = 7.6, p < .01, \eta_p^2 = 0.461$, see Fig. 9A. Post-hoc analyses revealed that all the three flight phases in real flight led to higher HR than their counterpart in VR ($p = 0.02$). In real flight setting, the take-off and the landing led to significantly higher HR than during downwind ($p < .001$). In VR setting, only the take-off induced significantly higher HR than during the downwind ($p = .02$).

Heart Rate Variability. A second two-way repeated ANOVA disclosed a main effect of the flight setting on RMSSD, $F(1, 9) = 17.2, p = .002, \eta_p^2 = 0.657$, as well as a main effect of the flight phase on heart rate, $F(2, 18) = 13.9$, $p = .001$, $\eta_p^2 = 0.608$, but no significant flight setting × phase interaction, see Fig. 9B. Post-hoc analyses revealed that all the RMSSD was lower during real flight than during VR flight ($p < 0.01$) and that HRV was lower during the landing than the two other phases ($p < 0.03$) and also lower during the landing compared to the take-off ($p < 0.03$).

Similar to RMSSD, a two-way repeated ANOVA on NN50 (Fig. 9C) revealed a significant main effect of flight settings, $F(1, 9) = 14.0, p = .005, \eta_p^2 = 0.608$, as well as a main effect of the flight phase, $F(2, 18) = 4.7, p = .042, \eta_p^2 = 0.345$. No significant interaction was found. Post-doc analyses showed lower NN50 values for the landing phase compared to the downwind ($p = .021$).

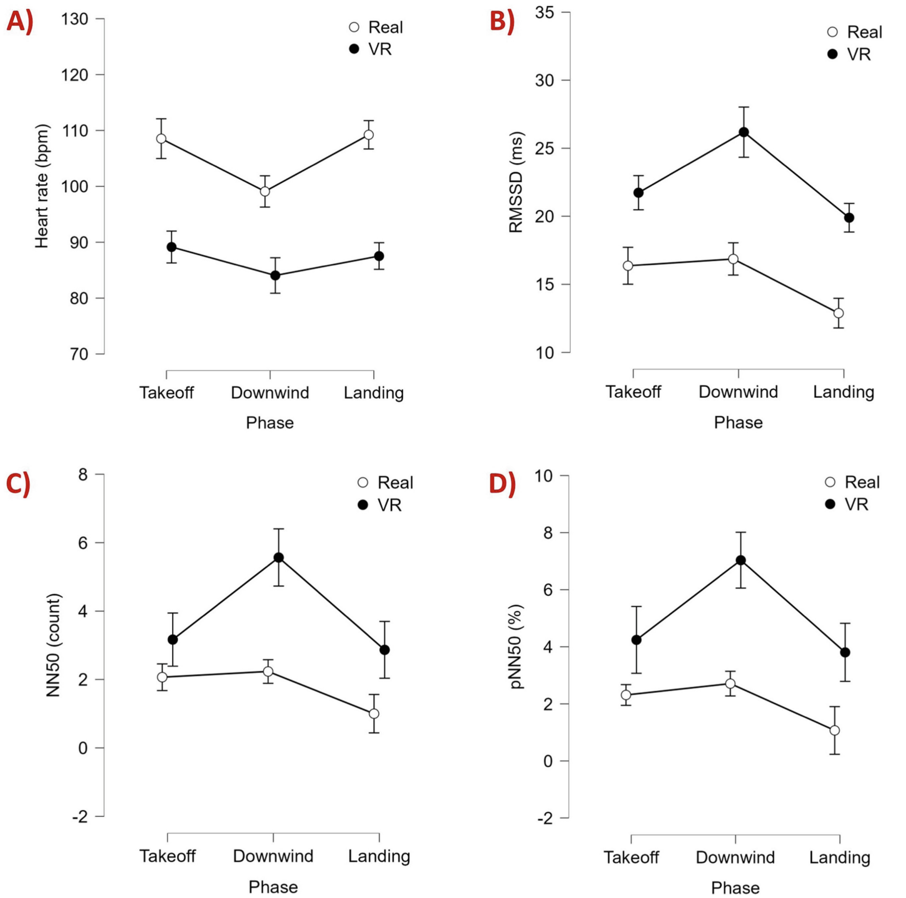

Fig. 9. A) mean HR (in bps); B) mean HRV (RMSSD in ms); C) mean NN50; D) mean pNN50 (%).

Finally, a two-way repeated ANOVA on pNN50 (Fig. 9D) revealed the same results as for the NN50 measure: a signiifcant main effect of flight settings, $F(1,9) = 14.5, p = .004, \eta_p^2 = 0.617$, as well as a main effect of the flight phase, $F(2,18) = 4.5, p = .049, \eta_p^2 = 0.332$. No significant interaction was found. Post-doc analyses showed lower NN50 values for the landing phase compared to the downwind ($p = .026$).

4 Discussion

In this study, we compared pilots' behavior through subjective questionnaires and cardiac activity during virtual reality and real flight. This study builds upon a previous conference paper [25] with a preliminary report of data from 4 subjects. In the present work, we analyze the data from 10 pilots who performed the same traffic pattern in a

DR400 light aircraft in a VR flight simulator and in actual flight condition. We also computed additional metrics of heart rate variability, i.e. NN50 and pNN50, as compared to only RMSSD in the previous study. Finally, using additional questionnaires allowed us to better understand what parts of virtual flight need to be considered in more detail to increase the realism of the simulation.

4.1 Subjective Measures

Sens of Presence. The analysis of the answers to the presence questionnaire revealed several positive aspects thus showing that the VR simulator was immersive. However, some negative points also appeared in these results such as a low definition of the graphics and interactions related issues.

Immersion. First of all, the feeling of immersion came out quite positively in the responses. Several factors contributed to this result: the quality of the global graphic environment, the similarity with a known real environment, and the realism of the audio environment of the simulation. The interactions with the aircraft also allowed the pilots to quickly feel in control.

The pilots, while performing this flight in an identically reproduced virtual environment, immediately felt at ease by judging themselves capable of controlling the events (question 1 = 5.17 $\pm$ 0.98). The score of question 9 shown a strong involvement of the pilots in the virtual environment experience (question 9 = 6.67 $\pm$ 0.52), this is even the highest score of the questionnaire. The other two highest scores were for questions 16 and 17 about the audio environment. Pilots were perfectly able to recognize the sounds (engine), which strongly invited them to be involved in the experiment (question 16 = 5.67 $\pm$ 1.03; question 17 = 5.67 $\pm$ 1.86).

They also felt involved because of some visual aspects of the environment (question 8 = 5.33 $\pm$ 1.75) and felt quite capable of anticipating the consequences of their actions (question 6 = 5.50 $\pm$ 1.38). In addition, the low score of question 10 represents a positive evaluation here, as this question allowed for judging the delay between actions and consequences (question 10 = 2.33 $\pm$ 1.03). Thus, the pilots did not feel any delay, which also contributed to their immersion in the experiment.

Finally, the score on the sensitivity of the environment following the pilots' actions (question 2 = 5.17 $\pm$ 0.98), and on the speed with which the pilots adapted to the virtual environment (question 1 = 5.17 $\pm$ 0.75) are also elements that suggest a good level of immersion.

All of these factors (interface quality, realism, and interactions) allowed the pilots to experience a strong sense of immersion, which fostered a high level of self-confidence, and allowed them to feel a strong sense of presence. This immersion did not generate any particular distraction since the maneuvers were perfectly performed by the pilots.

Points to Improve. Given the relatively low scores for some questions, several points of the virtual simulation can be improved. Question 3 disclosed that the interactions with the virtual environment are not necessarily well or poorly rated since the score is perfectly neutral on this scale from 1 to 7 (question 3 = 4 $\pm$ 1.26). The sensitive point was related to the absence of tracking of the hands in the virtual environment. While it was not an issue to operate the stick and the rudder since our participants could sense them with their hands and feet, it remained challenging to find and interact with the physical throttle and the flap levers. However, the technologies of motion capture and the synthesis of the virtual hands evolve quickly [21] including using own hands in mixed reality [11] for better seamless interactions, and while the current version of the VRtigo simulator did not include the hands' synthesis, the further studies will include these aspects.

Question 7 on the ability to visually explore the environment (question 7 = 4.17 ± 1.33) also highlighted a weakness noticed in a previous study [26]. The pilots also pointed out the difficulty to read accurately some critical cockpit instruments such as the anemometer due to relatively low graphics definition. This made the virtual flight more complicated than the real ones in this respect. However, todays' virtual reality headsets improved greatly since the beginning of this study, and recently released headsets such as Varjo XR-3 provide a higher resolution that could increase the readability of the instruments and parameter values.

Immersive Tendencies. Concerning the overall score of the ITQ, we know that the higher this score is, the more the participants will have a propensity to be immersed in the virtual reality experience. Here, the results show that this score is 79, which greatly exceeds the norms established by the Cyberpsychology laboratory at UQO (+15 points, i.e. +23%). Our group of participants, therefore, had a high propensity for immersion and the experiment could proceed.

The results by subscale show us first of all that the Focus subscale, even if slightly below the norm (24.67 vs. 24.81, i.e. –0.5%), corresponds well to it. Then, the Emotion subscale is slightly above the norm (15.75 vs 14.25, i.e. +10%). On the other hand, we note that the scores of the other two subscales are well above the norms (Involvement +39% and Game +77%). This could be explained by the relatively young age of the participants (22.4 ± 4.43 years), a generation more in contact with new technologies.

Flight Control in Virtual Reality. Keeping in the mind that in terms of virtual reality hardware on the VRtigo simulator, we only used the virtual reality headset and not the two classical stick controllers associated. We decided to provide the pilot with more intuitive aircraft control elements than virtual reality controllers (a rudder, a control stick, a flap lever, and a throttle). The answers to the questionnaire on the difficulties in accomplishing certain tasks or parts of the flight are therefore to be taken into consideration in relation to these control elements that are more typical of aircraft and not virtual reality controllers.

Questions 1 to 4 were related to the ability to control the aircraft (attitude, altitude, bank angle, and thrust), while questions 5 to 14 were related to the evaluation of the similarity of certain parts of the flight in real vs. virtual reality.

Control of the Aircraft. Regarding the first part of the questions about the controllability of the aircraft, our participants answered quite positively and felt comfortable with flying the virtual aircraft. Question 4, on the ease of controlling the thrust of the aircraft, done by the throttle lever, obtained the highest score on the questionnaire (score

= 6.5 ± 0.55). The other three scores, for attitude, altitude, and bank angle, were also highly rated with respective values of 5.83 ± 0.98, 5.67 ± 1.38, and 5.5 ± 1.21. These 4 scores are among the 5 highest on the questionnaire. This indicated that the pilots were comfortable with the different aircraft controls such as stick, rudder, and throttle. This observation on the controllability of the aircraft was important in our case of light aircraft piloting in virtual reality. This allowed pilots to feel involved and successful, as evidenced by the score of question 9 on involvement in the presence questionnaire (6.67 ± 0.52). These initial results corroborated from our previous 2018 study [26].

Similarity of the Virtual Experience. Responses to questions about the similarity of different parts of the traffic pattern between the virtual and real flights were more nuanced overall. Question 10, about the descending turn (score = 5.83 ± 0.75), and question 11 about the descent phase (5.50 ± 0.84) are the two best scores in this second part of the questions. The descending part of the traffic pattern is thus highly rated by the participants. However, we can also note that some phases get lower scores. Question 14 in particular, about the landing phase (score = 3 ± 1.55) as well as question 5 about the take-off phase (score = 4.33 ± 2.07). This shows us that the phases close to the ground, take-off, and landing, the most stressful and dangerous phases, in reality, were the least representative, which is also consistent with previous study [26].

One of the reasons could be related to some limitations of the virtual reality simulator already noted in the same study [26], in particular, the low resolution of the virtual reality headset, which negatively impacted the pilots' ability to correctly read the speed on the anemometer. We also found, during post-flight discussions with the pilots, that the physical sensations during the phases closed to the ground, mainly landing, were not completely realistic, even with the 6-axis mobile platform. The ground effects were not simulated realistically enough, which may have disturbed the pilots during the landing to execute flare phase. Another reason could be a lower experience level when using the VR simulator than the real aircraft.

4.2 Cardiac Activity

Heart Rate. As far as the real flights were concerned, the HR analyses showed that the closer the plane was to the ground (take-off and landing), the higher the HR was (Fig. 9A), with a maximum for landing (close to the take-off). The mean HR in real flight was 9.5% higher (take-off) and 10.2% higher (landing) than during the downwind phase. These results were consistent with previous findings reported during a traffic pattern experiment in real flight conditions [25, 37].

Regarding the virtual reality flights, it was interesting to note that the comparison of these three phases was similar to real flights with respectively an increase of 6.1% (take-off) and 4.1% (landing) compared to the downwind phase, with a maximum for takeoff.

In both environments (virtual and real), the HR of the downwind phase was, therefore, lower than that of the other two phases. We also noted, both in VR and in real flight, that the evolution of the HR following the three flight phases tended to be similar. The HR analysis also revealed a gap between virtual reality and real flights (Fig. 9A). The

HR was higher in real flight by about 22% (take-off), 18% (downwind), and 25% (landing) compared to virtual reality. This result could be interpreted as a lack of feeling of immersion experienced by participants in the VR condition. However, it is important to mention that the real flights were performed with crosswinds, especially for some pilots (16 G 26 kt at 30° from the runway axis). The crosswind induced in return higher mental demand (constant correction of trajectories). These aerological differences conditions could thus explain this difference between VR and reality findings.

Hear Rate Variability

The RMSSD Parameter. The analysis of the mean RMSSD during the real fight condition (Fig. 9B) disclosed higher values in downwind than during the take-off and landing phases, with a minimum for landing. These results for real flights were similar to previous studies [25, 37]. However, the difference of the RMSSD values between take-off and downwind was more pronounced in the 2021 study [25].

The results in virtual reality (Fig. 9B), followed a similar shape than in real flight with a higher average RMSSD in downwind than during the take-off and landing phases, and a minimum for landing too. Note however that the difference between take-off and downwind was much more marked than in virtual reality (–5 ms vs. –0.5 ms).

Again, these results seem to suggest that the real flight condition induced higher mental demand and psychological stress than the simulated condition [10]. The analysis of the RMSSD also revealed a gap between real flights and virtual reality. The RMSSD values were on average 25% (take-off), 36% (downwind), and 35% (landing) lower in reality than in VR. This is also in line with the results of the HR analysis, but with reversed lag.

The NN50 and pNN50 Parameters. The same analyses from the RMSSD were obtained for these two parameters describing the HRV (Figs. 9C and 9D). We have also a higher value for the downwind phase, and lower for the landing phase, for both virtual reality and real flights. Values for the three phases are also lower in real flights than in VR. The NN50 values were on average 35% (take-off), 60% (downwind), and 65% (landing) lower in reality than in VR. The pNN50 values were on average 46% (take-off), 62% (downwind), and 72% (landing) lower in reality than in VR.

4.3 Motion Sickness

The large majority of pilots did not report motion sickness after using virtual reality during the experiments. Only one pilot was slightly bothered. It must be said that the virtual flight lasted about 25 min, which can seem relatively long in terms of immersion for this kind of activity (important mental load for the management of a flight). This pilot immediately felt better once the experiment was over. This information this time confirms our previous study [25] for which we had too few participants.

4.4 Conclusion

Virtual reality technology continues to gain terrain in simulation and training with advanced state-of-the-art headsets with improved display resolution, embedded eye and hand tracking for natural interactions, and improved comfort. Recently, the first virtual flight simulator for rotary pilot training was qualified by European Aviation Safety Agengy. This first step of larger VR technology integration into the pilot training domain requires an in-depth understanding of how virtual flight differs from flying a real aircraft. While seeking the identical render of the flight instruments, as well as a motion platform for providing the pilot the vestibular cues, are paramount, neuroergonomic studies are also required [9, 15]. Such neuroergonomics evaluations allow to go beyond observable behaviors and compare different flight settings using objective neuro- and psycho-physiological measures such as electroencephalography, functional near-red spectroscopy, gaze tracking, or electrocardiography.

In this work, we compared standard flight patterns on a DR400 light aircraft in virtual reality (head-mounted device and motion platform) and real flight. We used subjective measures through different questionnaires of immersion, presence, ability to control the aircraft, but, most importantly, objective measures via cardiac activity. To the best of our knowledge, it is the first study comparing real and virtual flight (with a repeated measures analysis) using physiological measures.

The results showed that virtual flight is a decent resource for pilot training and pilots were able to fully perform the required actions for controlling the aircraft. Points to improve include updating the hardware (for better display resolution and hand tracking) but also to better model near-ground effects for the motion platform as these vestibular cues are important for the piloting experience. Heart rate and heart rate variability measures pointed out that real and virtual flights induce different stress and workload levels. The same pattern followed different flight phases in both virtual and real flights, i.e. more complicated and stressful operations during take-off and landing compared to the downwind phase yielding higher heart rates and lower HRV values. However, the overall heart rates were, generally, higher during the real experience (+20 bpm on average), and lower HRV values during the real flight compared to the virtual experience.

Overall, these results indicated that virtual flight represents real flight and can be used for pilot training. It reproduced the changes in stress/workload throughout the flight but a real flight is required to experience the level of cardiac activity associated with real operations. Further studies with simulators using more advanced hardware can further reduce these differences between the real and virtual flight experiences. However, replacing pilot training with exclusively virtual flight hours remains utopian at this point.

Acknowledgements. The authors would like to thank Stéphane Juaneda, the safety pilot, for his availability to perform flights and his precious know-how in-flight experimentation. Special thanks to Fabrice Bazelot and Benoît Momier, LFCL mechanics, for their help during the configuration of the experiments. A special thanks to Boris Jost and Alexandre Iche, ISAE-SUPAERO students, for their involvement in this project. Thanks to Guillaume Garrouste for the 3D development of the LFCL environment, and Jérôme Dartigues for building the mechanical part of the VRtigo platform.

References

1. Ahmed, S., Irshad, L., Demirel, H.O., Tumer, I.Y.: A comparison between virtual reality and digital human modeling for proactive ergonomic design. In: Duffy, V.G. (ed.) HCII 2019. LNCS, vol. 11581, pp. 3–21. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-22216-1_1
2. Auer, S., Gerken, J., Reiterer, H., Jetter, H.C.: Comparison between virtual reality and physical flight simulators for cockpit familiarization. In: Mensch und Computer 2021, pp. 378–392 (2021)
3. Ayoub, A., Pulijala, Y.: The application of virtual reality and augmented reality in oral & maxillofacial surgery. BMC Oral Health **19**(1), 1–8 (2019)
4. Bouchard, S., Robillard, G., Renaud, P.: Questionnaire sur la propension à l'immersion. Lab Cyberpsychologie L'UQO (2002)
5. Bric, J.D., Lumbard, D.C., Frelich, M.J., Gould, J.C.: Current state of virtual reality simulation in robotic surgery training: a review. Surg. Endosc. **30**(6), 2169–2178 (2016)
6. Chae, J.: Study on firefighting education and training applying virtual reality. Fire Sci. Eng. **32**(1), 108–115 (2018)
7. Clifford, R.M., Jung, S., Hoermann, S., Billinghurst, M., Lindeman, R.W.: Creating a stressful decision making environment for aerial firefighter training in virtual reality. In: 2019 IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 181–189. IEEE (2019)
8. Dehais, F.: Monitoring pilot's mental workload using erps and spectral power with a six-dry-electrode eeg system in real flight conditions. Sensors **19**(6), 1324 (2019)
9. Dehais, F., Karwowski, W., Ayaz, H.: Brain at work and in everyday life as the next frontier: grand field challenges for neuroergonomics. Front. Neuroergon. **1** (2020)
10. Durantin, G., Gagnon, J.F., Tremblay, S., Dehais, F.: Using near infrared spectroscopy and heart rate variability to detect mental overload. Behav. Brain Res. **259**, 16–23 (2014)
11. Feng, Q., Shum, H.P., Morishima, S.: Resolving occlusion for 3d object manipulation with hands in mixed reality. In: Proceedings of the 24th ACM Symposium on Virtual Reality Software and Technology, pp. 1–2 (2018)
12. Feng, Z., González, V.A., Amor, R., Lovreglio, R., Cabrera-Guerrero, G.: Immersive virtual reality serious games for evacuation training and research: a systematic literature review. Comput. Educ. **127**, 252–266 (2018)
13. Fussell, S.G., Truong, D.: Preliminary results of a study investigating aviation student's intentions to use virtual reality for flight training. Int. J. Aviat. Aeron. Aeros. **7**(3), 2 (2020)
14. Gateau, T., Ayaz, H., Dehais, F.: In silico vs. over the clouds: on-the-fly mental state estimation of aircraft pilots, using a functional near infrared spectroscopy based passive-bci. Front. Human Neurosci. **12**, 187 (2018)
15. Gramann, K., et al.: Grand field challenges for cognitive neuroergonomics in the coming decade. Front. Neuroergon. **2** (2021)
16. Grantcharov, T.P., Kristiansen, V.B., Bendix, J., Bardram, L., Rosenberg, J., Funch-Jensen, P.: Randomized clinical trial of virtual reality simulation for laparoscopic skills training. Brit. J. Surg. **91**(2), 146–150 (2004). https://doi.org/10.1002/bjs.4407
17. Hays, R.T., Jacobs, J.W., Prince, C., Salas, E.: Flight simulator training effectiveness: a meta-analysis. Milit. Psychol. **4**(2), 63–74 (1992)
18. Heathers, J.A.: Everything hertz: methodological issues in short-term frequency-domain hrv. Front. Physiol. **5**, 177 (2014)
19. Jensen, L., Konradsen, F.: A review of the use of virtual reality head-mounted displays in education and training. Educ. Inf. Technol. **23**(4), 1515–1529 (2018)
20. Joda, T., Gallucci, G., Wismeijer, D., Zitzmann, N.: Augmented and virtual reality in dental medicine: a systematic review. Comput. Biol. Med. **108**, 93–100 (2019)

21. Jörg, S., Ye, Y., Mueller, F., Neff, M., Zordan, V.: Virtual hands in vr: motion capture, synthesis, and perception. In: SIGGRAPH Asia 2020 Courses, pp. 1–32 (2020)
22. Kim, Y., Kim, H., Kim, Y.O.: Virtual reality and augmented reality in plastic surgery: a review. Arch. Plast. Surg. **44**(3), 179 (2017)
23. Kothe, C., Medine, D., Boulay, C., Grivich, M., Stenner, T.: Lab streaming layer (2014). https://github.com/sccn/labstreaminglayer
24. Koutitas, G., Smith, S., Lawrence, G.: Performance evaluation of ar/vr training technologies for ems first responders. Virt. Reality **25**(1), 83–94 (2021)
25. Labedan, P., Darodes-De-Tailly, N., Dehais, F., Peysakhovich, V.: Virtual reality for pilot training: study of cardiac activity. In: VISIGRAPP (2: HUCAPP), pp. 81–88 (2021)
26. Labedan, P., Dehais, F., Peysakhovich, V.: Evaluation de l'expérience de pilotage d'un avion léger en réalité virtuelle. In: ERGO'IA (2018)
27. Lawrynczyk, A.: Exploring Virtual Reality Flight Training as a Viable Alternative to Traditional Simulator Flight Training. Ph.D. thesis, Carleton University (2018). https://doi.org/10.22215/etd/2018-13301
28. Li, L., et al.: Application of virtual reality technology in clinical medicine. Am. J. Transl. Res. **9**(9), 3867 (2017)
29. Markopoulos, E., et al.: Neural network driven eye tracking metrics and data visualization in metaverse and virtual reality maritime safety training (2021)
30. Meshkati, N.: Heart rate variability and mental workload assessment. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload, Advances in Psychology, vol. 52, pp. 101–115. North-Holland (1988). https://doi.org/10.1016/S0166-4115(08)62384-5, http://www.sciencedirect.com/science/article/pii/S0166411508623845
31. Oberhauser, M., Dreyer, D., Braunstingl, R., Koglbauer, I.: What's real about virtual reality flight simulation? Aviation Psychology and Applied Human Factors (2018)
32. Peysakhovich, V., Monnier, L., Gornet, M., Juaneda, S.: Virtual reality versus real-life training to learn checklists for light aircraft. In: 1st International Workshop on Eye-Tracking in Aviation (2020)
33. Pottle, J.: Virtual reality and the transformation of medical education. Fut. Healthcare J. **6**(3), 181 (2019)
34. Pourmand, A., Davis, S., Lee, D., Barber, S., Sikka, N.: Emerging utility of virtual reality as a multidisciplinary tool in clinical medicine. Games Health J. **6**(5), 263–270 (2017)
35. Galasko, C.S.: Competencies required to be a competent surgeon. Ann. Roy. Coll. Surg. Engl. **82**, 89–90 (2000)
36. Sakib, M.N., Chaspari, T., Behzadan, A.H.: Physiological data models to understand the effectiveness of drone operation training in immersive virtual reality. J. Comput. Civil Eng. **35**(1), 04020053 (2021)
37. Scannella, S., Peysakhovich, V., Ehrig, F., Lepron, E., Dehais, F.: Assessment of ocular and physiological metrics to discriminate flight phases in real light aircraft. Hum. Fact.: J. Hum. Fact. Ergon. Soc. **60**(7), 922–935 (2018). https://doi.org/10.1177/0018720818787135
38. Silverstein, J.C., Dech, F., Edison, M., Jurek, P., Helton, W., Espat, N.: Virtual reality: immersive hepatic surgery educational environment. Surgery **132**(2), 274–277 (2002). https://doi.org/10.1067/msy.2002.125723, http://www.sciencedirect.com/science/article/pii/S0039606002000843
39. Tarvainen, M.P., Niskanen, J.P., Lipponen, J.A., Ranta-Aho, P.O., Karjalainen, P.A.: Kubios hrv-heart rate variability analysis software. Comput. Methods Progr. Biomed. **113**(1), 210–220 (2014)
40. Thomsen, A.S.S., et al.: Operating room performance improves after proficiency-based virtual reality cataract surgery training. Ophthalmology **124**(4), 524–531 (2017)
41. Togo, F., Takahashi, M.: Heart rate variability in occupational health ___a systematic review. Ind. Health **47**(6), 589–602 (2009). https://doi.org/10.2486/indhealth.47.589

42. Walters, W.T., Walton, J.: Efficacy of virtual reality training for pilots: a review of links between user presence, search task performance, and collaboration within virtual reality. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, vol. 65, pp. 919–922. SAGE Publications, Los Angeles (2021)
43. Witmer, B.G., Singer, M.J.: Measuring presence in virtual environments: a presence questionnaire. Presence **7**(3), 225–240 (1998)
44. Xie, B., et al.: A review on virtual reality skill training applications. Front. Virt. Reality **2**, 49 (2021)
45. Zhang, H.: Head-mounted display-based intuitive virtual reality training system for the mining industry. Int. J. Min. Sci. Technol. **27**(4), 717–722 (2017)