Behavioral and Physiological Assessment of a Virtual Reality Version of the MATB-II Task

Zoe Gozzi[1], Vsevolod Peysakhovich[2][a], Alma Cantu[2,3][b] and Mickael Causse[2][c]

[1] *Labsoft, Toulouse, France*

[2] *ISAE-SUPAERO, Université de Toulouse, France*

[3] *School of Computing, Newcastle University, U.K.*

Keywords: MATB-II, Mental Workload, Virtual Reality, Human Factors, Aviation.

Abstract: The goal of this research was to examine the possible benefits of adapting the Multi-Attribute Task Battery (MATB-II) in a virtual reality (VR) environment to provide an immersive and ecological platform for studies on mental workload in the aerospace domain. The original desktop MATB-II has many advantages, but the level of immersion remains moderate, and the computer screen greatly reduces the spatial dimension existing in real environments such as the cockpit. Thirty-one participants performed an experiment during which we compared the original MATB-II with the new virtual version, called "MATB-II VR". We used subjective, performance, and cardiovascular measurements. The virtual MATB-II was performed without ("MATB-II VR No Touch") and with tactile feedback ("MATB-II VR Touch"). In general, the results showed that mental and physical efforts were higher and performances lower with the virtual version. Heart rate was higher with the virtual version, supporting the idea that such environment is more challenging. The individual performance in the desktop and the virtual environments correlated well, showing that our virtual version engaged analog physical and cognitive abilities as compared with the original version. Interestingly, performance during MATB-II VR was well predicted by basic mental rotation performance assessed with a neuropsychological task.

1 INTRODUCTION

MATB-II (Multi-Attribute Task Battery II) is a computer-based task (Comstock & Arnegard, 1992) that can be used to evaluate operator performance and workload by providing a set of four concurrent subtasks (Santiago-Espada et al., 2021), analogous to ones that aircrews perform in flight. The wide usage of MATB-II allows comparing results across a great variety of studies. Indeed, more than 135 research papers have been published using the MATB-II as an experimental platform, for example in aerospace medicine (Chandra et al.,2015), psychology (Daviaux et al., 2019), human factors (Kennedy et al., 2017), or alarm system design (Chancey et al., 2015). In the context of aeronautical research, 2D synthetic tasks such as MATB-II has many advantages, including flexibility, cost-efficiency, and the fact that they do not require complex hardware to be implemented. On the other hand, their level of immersion remains relatively moderate, and tasks performed on a computer screen greatly reduce the spatial dimension existing in real environments. Yet, the spatialization of the instruments and the multiple interactions with various flight controls in the cockpit are an important aspect of piloting (Letondal et al., 2018), and thus visual spatial attention is highly solicited. Virtual reality (VR) offers a solution halfway between desktop simulations and a full flight simulator. Increasing immersion may be useful to reproduce the three-dimensional space of the cockpit (Oberhauser et al. 2015). VR settings can create reproducible environments for real-time experiments, where the participants can sense, feel, and interact with the virtual world. The technology incorporates visual and acoustic stimuli and creates an immersive and ecological situation that allows being isolated from distractions that could bias experiments. The entire body may be implicated in the interaction with VR, and the participants can interact

[a] https://orcid.org/0000-0002-9791-4460

[b] https://orcid.org/0000-0001-6081-2439

[c] https://orcid.org/0000-0002-0601-2518

Gozzi, Z., Peysakhovich, V., Cantu, A. and Causse, M.
Behavioral and Physiological Assessment of a Virtual Reality Version of the MATB-II Task.
DOI: 10.5220/0010912100003124
In *Proceedings of the 17th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2022) - Volume 2: HUCAPP*, pages 77-87
ISBN: 978-989-758-555-5; ISSN: 2184-4321

with the virtual content not only through a controller but also directly with their hands. Not surprisingly, VR is extensively used as a research tool in different professional domains like military training (Lele, 2013), medical training (Falah et al., 2014; Singh et al., 2020; Smith et al., 2020), or teaching (Sympnenko et al., 2020). In the domain of the human factors in aviation, it provides an easy way to reproduce complex piloting environments and simulated piloting scenarios (Labedan et al., 2021; Peysakhovich et al., 2020).

Given the lower immersion of 2D environments, the use of virtual reality to reproduce life tasks and situations has grown rapidly (Hassandra et al., 2018; Ansado et al., 2021; Zygouris et al., 2017; Soret et al., 2019). Numerous studies have focused on the differences in task performance between VR and Desktop environments, and the results are sometimes inconsistent. Pausch (1997) showed that users with a VR interface complete a search task faster than users with a desktop display (the desktop display was implemented in the VR headset to create a "stationary monitor"). The authors also found a positive transfer of training from VR to desktop display and a negative transfer of training from desktop displays to VR (Pausch et al., 1997). Pallavicini's (2019) research did not indicate any differences between video games played in VR and with a desktop in terms of usability and performance. Whereas the authors mention that previous literature reported better performances in non-immersive display modalities due to better usability, the researchers hypothesized that with the progress in technology, this difference will no longer exist. In addition, researchers found that VR enhances emotional arousal thanks to the "wow-effect", the temporary state that new technology triggers in individuals when they are exposed to a new experience. In aeronautics, Oberhauser (2018) investigated the functional fidelity of a virtual reality flight simulator in comparison with a conventional flight simulator. Their results showed that the deviations in flight performance (heading, altitude, flight path deviations, delays in operating the controls) were significantly larger in VR than in the conventional flight simulation. Yet, most participants could safely and reliably complete the flight task. Besides, the pilots reported a higher workload in the virtual environment. Without substituting to real simulators, VR could be an interesting and viable tool to perform human factors research that reproduces immersive, engaging, and complex tasks such as the MATB-II.

The goal of this exploratory study was to assess the added value of implementing a task close to the original 2D version of the MATB-II in VR for aerospace research. We believe that a VR version of the MATB-II can provide a more ecological platform to perform behavioral and physiological measurements (Luong et al., 2020) and can solicit more visuospatial skills (Maneuvrier et al., 2020) than the desktop version. Moreover, the VR environment will be a closer simulation of a cockpit environment. We compared the 2D and 3D versions, in particular with respect to mental and physical effort, engagement, and their ability to engage particular cognitive abilities such as multitasking and visuospatial skills. Thirty-one participants performed the MATB-II in two different environments: the original MATB-II performed on a desktop computer and the new version of the MATB-II performed in immersive conditions and called the "MATB-II VR". Participants also performed three neuropsychological tasks, a multitasking task, a mental rotation task, and a visual search task. Participants' outcomes to these tasks were compared with performances obtained in the 2D and 3D versions of the MATB-II. All along with the MATB-II and MATB VR task performance, electrocardiogram (ECG) measurements were performed to better characterize the level of mental effort of the participant (Kim et al., 2018), in relation to mental workload and task engagement.

We hypothesized that the MATB-II VR should be more engaging, elicit a higher workload, and performance inside this environment should be better predicted by the result obtained with the neuropsychological tests, in particular the visuospatial ones, since VR environments may solicit more theses functions. We also hypothesized that heart rate should be higher in the MATB VR versions versus the desktop MATB-II version since the VR environment create a more stressful situation. We had no clear hypothesis regarding task performance between the desktop and the immersive environment due to the common issues when interacting inside a VR environment (Geszten et al., 2018), especially because the MATB-II task requires interacting with relatively small elements. In order to better address this question, the MATB VR was performed two times, one time without a particular device and another time with the GO VR Touch device. The GO VR Touch was placed on the left index and aimed at reproducing the haptic sensation in the VR environment, in particular when touching the buttons. When the task was performed with this device, it was called "MATB VR Touch". We hypothesized that performance could be better in the MATB VR Touch vs MATB VR No Touch.

2 METHOD

2.1 Participants

31 participants (age = 31.2 ± 9.2 years; 19 male, 12 female) took part in the experiment. They were recruited among students and employees. Thirty participants were right-handed. 2 participants only had pilot experience but they were all knowledgeable in the field of aeronautic. They all had a sufficient level of English in order to understand the instructions and the audio communication during the MATB-II tasks. One participant reported a beginner level of English, 13 intermediary levels, 16 high proficiency, and 1 native speaker. The participants did not report any heart-related problems. 17 participants reported to be ”not at all fatigued”, 11 participants reported to be ”somewhat fatigued” and 3 participants reported to be ”very fatigued”.

2.2 MATB-II Original Task

The original desktop MATB-II is made of four separate subtasks: the system monitoring (SYSMON) subtask, the tracking (TRACK) subtask, the communication (COMM) subtask, and the resource management (RESMAN) subtask, see Fig. 1. The SYSMON subtask is presented on the top left corner of the screen (number 1, Fig. 1). It simulates the monitoring of gauges and warning lights. The participant has to click with a mouse on one of the six items if an abnormal state occurs. In particular, in case of absence of the green light, presence of the red light, and if one of the four moving pointers deviates from the midpoint. Performance is measured by the reaction time to indicate an abnormal behavior. The TRACK subtask is presented on the top center of the screen (number 2, Fig. 1). It simulates manual aircraft control. Using the joystick, the participant has to keep the target at the center of the window, which is not easy since it is moving continuously in an erratic manner. Performance is generally measured by the average distance of the target to the center of the window. The COMM subtask is presented on the bottom left corner of the screen (number 3, Fig. 1). It simulates the management of air traffic control communications. This subtask presents pre-recorded auditory messages to the operator at selected intervals. However, not all messages are relevant to the operator. The participant’s task is to identify the ones that are relevant (according to a particular aircraft called sign) and to respond by selecting the appropriate radio and frequency on the COMM window with repetitive mouse clicks on the buttons. No action is required for messages with other call signs than the one attributed to the participant. Performance is characterized by the percentage of correct responses. The RESMAN is presented on the bottom center of the screen (number 4, Fig. 1). It simulates fuel management. The six large rectangular regions are tanks that hold fuel. The green levels within the tanks represent the amount of fuel in each tank, and these levels increase and decrease as the amount of fuel in a tank changes. The goal is to maintain tanks A and B at 2500 units each. This is done by turning On or Off the eight pumps with mouse clicks. Pump failures can occur and are shown by a red area on the failed pump. The performance is measured by the difference between the average fuel amount in each tank during the session vs the target amount (2500 units). The MATB-II task duration was 10 minutes, and it was entirely performed with the mouse and a joystick.

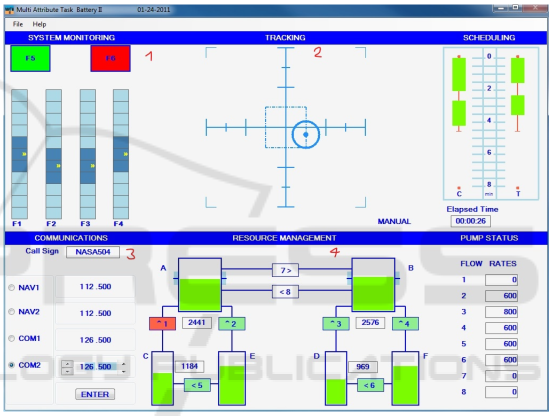

Figure 1: Interface of the original desktop MATB-II.

2.3 MATB-II VR Task

We developed the **MATB VR**, a task inspired from the MATB-II and adapted to benefit the advantages of the virtual reality environment Fig. 3. It integrates the same four subtasks as the original MATB-II. They have been spatially distributed to resemble a real cockpit. For example, the TRACKING subtask is located approximately where the artificial horizon is displayed in a real cockpit. Similarly, the SYSMON subtask is located approximately where the pilots display the fuel quantity. Some subtasks were slightly modified in comparison to the original MATB-II. This was done to better fit with the VR constraints, in particular regarding the relative difficulty to interact with objects. The COMM subtask had two channels instead of four and the SYSMON subtask had two scales instead of four. The MATB-II VR was performed two times, either with the GO Touch VR, this variant was called the **MATB-II VR Touch**, or without the GO Touch VR, this variant was called **MATB VR No Touch**. When referring to the virtual MATB-II task in general (without considering the presence or absence of the Go Touch VR), we simply used the term **MATB VR**. The MATB-II VR task duration was 5 minutes. The participant held the joystick with the right hand (TRACKING subtask), like during the original MATB-II, and performed the three other subtasks with the left hand, see Fig. 2.

Figure 2: The experimental setup. The participant is wearing an HTC Vive headset equipped with the Leap Motion Controller. In her left hand, she is wearing the Go Touch VR controller. She holds a joystick Cyborg X Flight Stick in her right hand to perform the TRACK task.

2.4 Calibration of the Difficulty during the MATB-II Tasks

We used the baud rate formula to set the difficulty of the two MATB-II environments. The baud rate is used to obtain a standard evaluation of the workload in a task (quantity of information over time) (Liu, 2018). In this way, task designers can qualify tasks as having low, medium, or high workloads by applying the baud formula. This allows for easy replication among different studies that use the same task and aim to control the difficulty and workload more efficiently. It is based on the Information theory of Shannon (1948) and it is defined as the number of possible tasks in bits, divided by the time in seconds (Camnden, 2017).

The formula to calculate the baud rate is as follows:

$$B(i) = \frac{H(i)}{\Delta T(i)} \quad (1)$$

In order to make the two environments (MATB-II and the MATB VR) relatively comparable in difficulty and feasible, we set the same baud rate, as ”low” across both of them for the SYSMON, COMM, and RESMAN, subtasks. A low baud rate is B(i) < 0.2, where ’i’ was SYSMON, COMM, and RESMAN subtasks. More precisely, it resulted in having 1 stimulus every 20 seconds in the SYSMON subtask, 7 audio communications every 5 minutes in the COMM subtask, and 1 stimulus every 40 seconds in the RESMAN subtask. We could not easily make the TRACK subtask comparable in the MATB-II and MATB-II VR, in particular because the erratic movement of the target was difficult to reproduce. One performance measure was considered per subtask: the number of correct responses during COMM subtask, the difference between the optimal level of the tank A and B (2500 units) and the fluctuations of the level of fuel during RESMAN, and the RMSD of the target from the center of the window during TRACK.

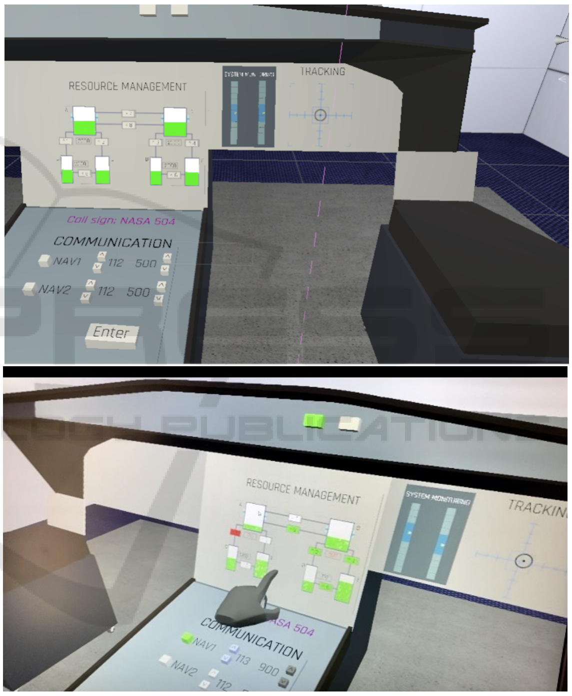

Figure 3: Interface of the MATB-II VR.

2.5 Neuropsychological Tasks

Three computer-based neuropsychological tests from the PsyToolkit (Stoet, 2010,2017) were performed by the participants. The results from these three tests were correlated with participants’ performances to the desktop MATB-II, the MATB VR No Touch, and the MATB VR Touch.

The multitasking task evaluates the ability to switch between two tasks. The participant is presented with two types of shapes, diamonds, or rectangles, filled with 2 or 3 dots. In the ”shape” condition, the participant has to press ”b” if the shape is a diamond and “n” if the shape is a rectangle. The participant has to ignore the dots. In the ”filling” condition, the participant has to press “n” if two dots fill the shape, and “b” if three dots fill the shape. The participant has to ignore the outer shape here. The rule to follow (shape or filling) depends on the location of the stimuli, on the top of the screen for shape, on the bottom of the screen for filling. Participants performed the first block, corresponding to a training period with 6 stimuli for the shape task and 6 stimuli for the filling task. After the training, participants performed three experimental blocks in the same order, with respectively 20 stimuli for the filling task, 20 stimuli for the shape task, and 40 stimuli in a condition where filling and shape alternate. The interval between two stimuli was 1 second. We measured the reaction time of the participants, see Fig. 4.

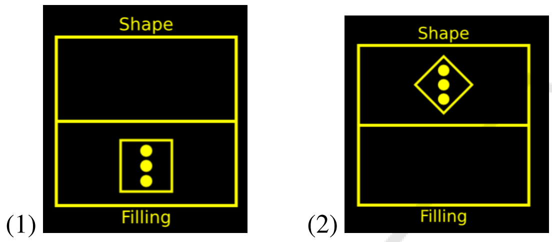

Figure 4: Multitasking task. Filling condition (1) and shape condition (2). A computer-based tasks designed with the use of PsyToolkit (Stoet, 2010;2017).

The mental rotation task evaluates the capacity to imagine what a stimulus would look like if it would be rotated. The participant is presented with three objects like in the example Figure 5. The participant has to decide which one of the bottom two matches the one on the top after rotations. The mental rotation task is a good predictor for visuospatial abilities and is used in aeronautical training and test batteries for the pre-assessment of pilot candidates (Krüger et al., 2016; Sladky et al., 2016). The task consists of a training block with 5 stimuli and then an experimental block with 10 stimuli. The participants had 20 seconds to respond. The tasks took about 2 minutes to be completed. We measured the reaction time of the participants.

The visual search task evaluates the ability to find a target stimulus among distractors. The participant is presented with 5, 10, 15, or 20 items, consisting in orange and blue letters ”T”. Blue T (always presented upward) and downward orange T are the distractors, and must be ignored. The participant has to press the space bar when an orange and regular upright position letter “T” is displayed on the screen. Participants has 4 seconds to respond, there were 50 search displays, and the task takes around 5 minutes to complete. Search time usually increases with large numbers of items on the screen. We measured the reaction time of the participant, see Fig. 6.

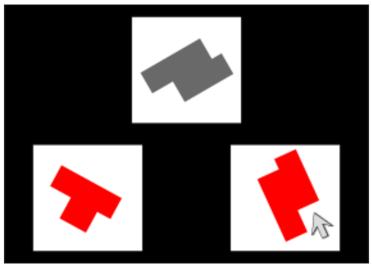

Figure 5: Mental Rotation Task. A computer-based tasks designed with the use of PsyToolkit (Stoet, 2010;2017).

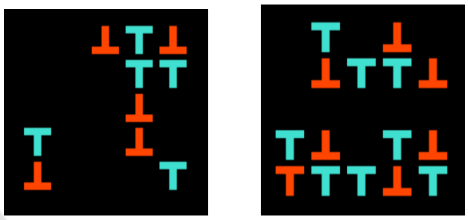

Figure 6: Visual Search Task. A computer-based tasks designed with the use of PsyToolkit (Stoet, 2010,2017).

2.6 Procedure

First, participants filled out a consent form and were then equipped with the ECG. They perform the three MATB-II tasks (MATB-II, MATB VR No-Touch, MATB VR Touch) and the three neuropsychological tasks in pseudo-random order (MATB-II and neuropsychological tasks were not mixed). After the performance of each MATB-II task, participants had to evaluate their level of mental effort, physical effort, and engagement using a 7 points scale, where 1 was the minimum level and 7 was the highest level. At the end of the experiment, the participant filled out a questionnaire to specify their familiarity with the virtual reality technology, their sleep habits, and different demographic information (age, gender,...). The total experiment duration was approximately 1 hour.

2.7 Experimental Material

Virtual Reality Headset and Virtual Environment. We used an HTC Vive head-mounted display (1080×1200 pixels per eye, 90 Hz, 110 degrees field of view). The Unity 3D engine and the C# programming language were used to develop the MATB-II VR. An optical hand-tracking LEAP Motion controller was physically mounted to the front of the VR headset to allow natural finger motions as input for the MATB-II VR environment.

Haptic Device. In the MATB-II VR Touch variant, the Go Touch VR controller was placed on the left index of the participant to reproduce haptic feedback when interacting with the buttons. The Go VR Touch controller creates pressure on the fingertip when interacting with the objects. The participants wore a Go Touch VR controller on the left index.

Joystick. A Cyborg X Flight Stick was used to perform the TRACK subtask in all versions of the MATB-II.

Electrocardiogram Measurements. Heart rate (in BPM) was measured with an ECG, using 3 electrodes placed on the thorax. The BIOPAC MP150 (System Inc, Santa Barbara, CA) software was used for the acquisition of the signal. The ECG signal was sampled at 1000 Hz and recorded using the AcqKnowledge software (BIOPAC System Inc, Santa Barbara, CA). The mean heart rate during each MATB-II variant was computed with the Kubios® software.

2.8 Statistical Analysis

Behavioral performance was processed using a home-build python script and all data was analyzed with R statistical software. As the data was not normally distributed, in particular subjective evaluations, we used the Wilcoxon signed-rank test to compare the conditions. Regarding performances comparisons across the three MATB-II variants, we focused the analysis on the COMM and RESMAN subtasks. We also conducted correlation and linear regression analyses using the performance measures of each task (MATB-II and neuropsychological tasks). For these correlations, we used TRACK, RESMAN, and COMM subtasks. SYSMON performance was not used to reduce the number of variables.

3 RESULTS

Some participants reported difficulties when equipped with the Go Touch VR controller. Also, some participants reported a delay between the action and the result while pushing some buttons in the virtual environment. These issues are discussed in more detail in the limitations section.

3.1 Subjective Results

Mental Effort. Participants reported a significant higher mental effort with the MATB-II VR Touch vs both the MATB-II VR No Touch (V = 71.5, p = .049) and the desktop MATB-II (V = 54, p = .023), see Fig. 7. The difference between the desktop MATB-II and the MATB-II VR No touch was not significant ($p > .05$).

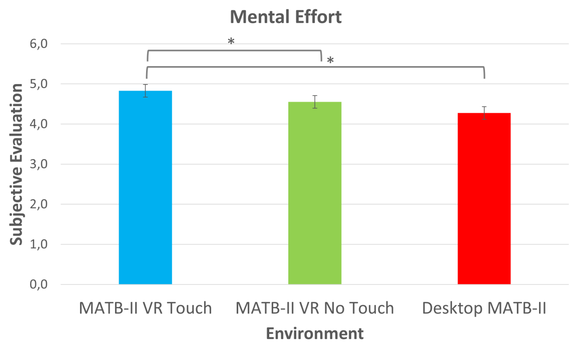

Figure 7: Subjective mental effort assessment across the three MATB-II variants.

Physical Effort. Participants reported a higher subjective physical effort during the MATB-II VR Touch vs the desktop MATB-II (V = 192, p = .029). The other comparisons were not significant (ps $>$ .05), see Fig. 8.

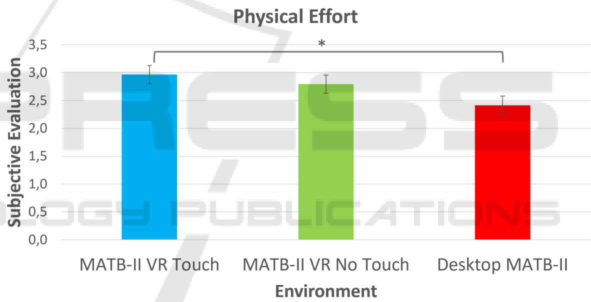

Figure 8: Subjective physical effort assessment across the three MATB-II variants.

Engagement. We found no difference across the three MATB-II task variants regarding the level of engagement (all $ps > .05$), see Fig. 9.

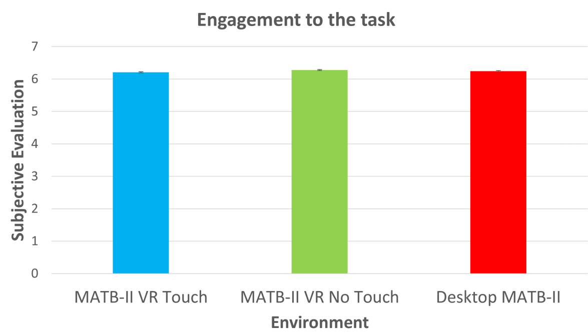

Figure 9: Subjective engagement assessment across the three MATB-II variants.

3.2 Performances Results

COMM Subtask. The number of correct responses was 90.8% in the desktop MATB-II, 86.6% in the MATB-II VR Touch, and 84.3% in the MATB VR No Touch, Fig. 10. The difference was significant between the desktop MATB-II and the MATB-II VR No Touch version (V = 72, p = .044). As a complementary analysis, we investigated the time taken to set the radio and the frequency. We observed a general longer completion time in the MATB-II VR environment than in the desktop MATB-II (3.60 ± 2.50 s). This is most probably due to some usability issues with the VR environment and the different interactions (clicking a mouse versus pushing down a virtual button). Interestingly, the completion time in the MATB-VR Touch (11.86 ± 6.54 s) was faster than in the MATB-VR No Touch (13.02 ± 8.34 s), possibly due to the haptic functionality giving feedback after every correct action.

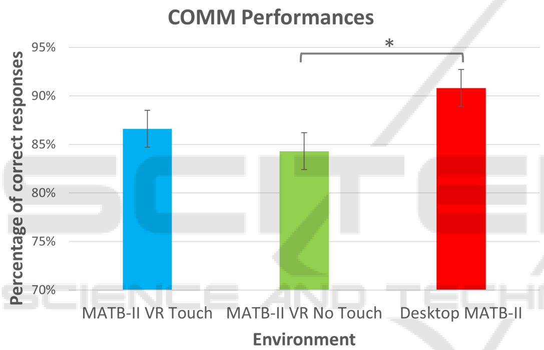

Figure 10: Performances at the COMM subtask across the three MATB-II variants.

RESMAN Subtask. We did not find any significant difference in RESMAN performance across the three MATB-II variants ($p > .05$), Fig. 11.

3.3 Electrocardiogram Results

ECG data were rejected for two participants due to recording issues. We found a significantly higher heart rate during MATB-II VR No Touch vs the desktop MATB-II (V = 310, p-value = .045). The other comparisons were not significant ($p > .05$), Fig. 12.

3.4 Correlation Analyses

A correlation analysis was conducted using all performance variables from the three MATB-II variants and the three neuropsychological tasks. The correlations shown in the Fig. 13 are significant at $p < .05$. The results indicated that the MATB-II transposed well in

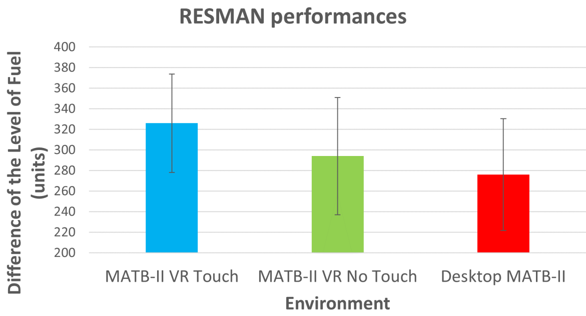

Figure 11: Performances at the RESMAN subtask across the three MATB-II variants.

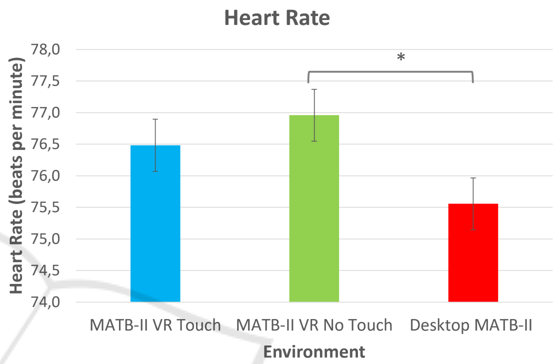

Figure 12: Heart Rate (beats per minutes) across the three MATB-II variants.

the VR environment since performance with the desktop and VR versions were highly correlated. This was the case for the three analyzed tasks: the TRACK, the RESMAN, and the COMM subtasks, see Fig 10.

Regarding the neuropsychological tasks, the performance during the mental rotation task correlated with the RESMAN subtask of the desktop MATB-II (r(30) = 0.56, $p < .05$), but this correlation was stronger with the RESMAN subtask of the MATB-II VR Touch variant (r(30 = 0.81, $p < .05$).

The mental rotation task performance also correlated with the TRACK subtask of both the desktop MATB-II (r(30) = 0.57, $p < .05$) and the MATB-II VR No Touch (r(30 = 0.47, $p < .05$). The mental rotation task performance was also positively correlated with the COMM subtask performance in the MATB-II VR No Touch (r(30) = -0.52, $p < .05$), while it was negatively correlated with the COMM subtask performance of the MATB-II VR Touch variant. The multitasking and the visual search tasks did not correlate significantly with the performance during the MATB-II tasks.

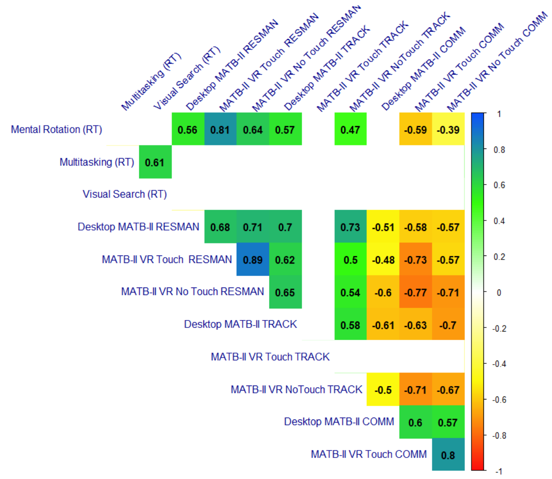

Figure 13: Correlation Matrix of the performances in the three MATB-II variants and during the three neuropsychological tasks. Only significant results are displayed.

3.5 Linear Regression

Based on the observed correlations between the mental rotation task and the RESMAN subtask during the three MATB-II variants, three simple linear regressions were performed to evaluate the possibility to predict (causal relationship) performance in the RESMAN subtask with the mental rotation task performance. Significant results were found for all three variants, the desktop MATB-II, ($F(1,29)$ = 13.19, = $p < .001$), $R^2 = 0.31$), the MATB-II VR No Touch, ($F(1,29) = 20.33$, $p < .001$, $R^2 = 0.41$), and the MATB-II VR Touch ($F(1,29) = 53.59$, $p < .001$), $R^2 = 0.64$). To summarize high performance in the mental rotation task predicted high performance in the RESMAN subtask and this relationship was higher in the VR environment, in particular during the MATB-II VR Touch variant.

4 DISCUSSION

In this study, we implemented the original desktop version of the MATB-II in virtual reality in an attempt to provide an immersive and ecological platform for studies in aerospace. We measured the performance and the cardiovascular activity during the original MATB-II version and our virtual variant, the latter was performed with or without tactile feedback. We also examined the relationship between the performances attained during the MATB-II variants with the performance obtained during three neuropsychological tasks, in order to better understand which cognitive abilities are engaged during each MATB-II variant. As expected, the results showed that participants evaluated the MATB-II VR environment as more mentally and physically demanding, in particular when performed with haptic feedback. These results are compatible with ones observed by Oberhauser (2018) in the context of flight simulation. The VR environment likely required more mental effort since participants had to pay more attention to the different tasks because stimuli did not always appear in the field of view, contrary to the desktop MATB-II (Wismer et al., 2021). In addition, the more important physical effort can be related to the extensive engagement of the body in the task, in particular with physical actions on the “virtual cockpit”. The fact that this effect was significant only with the tactile feedback is consistent with some reports from the participants after the experiment. Some of them reported that the Go Touch VR controller sometimes provided feedback with a time delay after pushing the buttons or was sometimes not functioning. Some participants also reported that they were bothered about the vibration due to the haptic feedback. These issues probably contributed to increase mental and physical efforts but did not necessarily degrade performance. On the contrary, while performance in the COMM subtask was lower in VR when no haptic feedback was provided (i.e., MATB-II VR No Touch), performance did not differ between the desktop version and MATB-II Touch, during which the Go Touch VR controller was used. In addition, in MATB-II VR, completion times to the COMM subtask were also better (shorter) when haptic feedback was provided. Despite some remaining issues, haptic feedback seems promising to improve interaction in VR.

In VR, the COMM subtask was one of the most difficult because it required interacting with small buttons with the finger, which strongly engaged participants’ attentional resources. It seems that interacting with the small buttons was even harder without the tactile feedback due to the above-mentioned issues. Interestingly, the correlation analysis showed that the COMM performances mostly correlated negatively with all the other tasks. In fact, it means that better performance to the COMM subtask was done at the expense of all other tasks. We could sometimes observe some participants disengaging from the TRACK subtask, stopping acting on the joystick, while managing the COMM subtask.

Cardiovascular results were consistent with the idea that the VR environment is more demanding since the heart rate was higher during the MATB-II VR No touch vs the desktop MATB-II. The difference between the MATB-II VR Touch vs desktop MATB-II was not significant. We can only speculate that this effect would have been significant with more statistical power. The correlation analyses performed on the performance during the three different MATB-II variants suggest a good transposition of the task from the desktop environment to the virtual one. All investigated subtasks correlated across the desktop MATB-II and the MATB-II VR. In other words, a participant that demonstrated high performance during RESMAN or COMM in the desktop MATB-II was also very likely to have high performance to the same subtasks in the virtual environment. This latter result suggests that the MATB-II VR can be a genuine alternative to the desktop MATB-II for immersive aerospace experiments. Finally, we found that reaction time during the mental rotation task was a good predictor of the performances attained during the RESMAN subtask of the desktop and VR MATB-II variants, explaining up to 64% of the performance variation.

5 LIMITATIONS AND REMARKS

This study has some limitations mainly due to the well-known usability problems in VR that were similarly observed in different studies like in Pallavicini (2019) and Santos (2009). Some feedback received from participants was about the technical issues regarding the VR material that sometimes increased frustration and stress. For example, the LEAP motion controller could regularly lose the tracking for some participants, requiring them to move their hand in front of the controller to recover the tracking. This has certainly contributed to jeopardizing some results in the VR environment. This relative difficulty to interact in the VR environment was also observed in the COMM subtask, with a longer time taken to select the radio and the frequency in the VR in comparison with the desktop environment. Additionally, participants had some problems with the Go Touch VR controller, which did not always work as expected. The haptic feedback was not always synchronized with the button press. In these cases, some participants reported that they did not know if the button press was taken into account because they expected synchronized visual and haptic feedback. This has led some participants to push the button several times. Despite this issue, still in the COMM subtask, we observed a shorter time taken to select the radio and the frequency when the Go Touch VR controller was used. Finally, some participants also reported that the VR headset became heavy and hot. This could be a significant issue for very long experiments as physical pain or unease could bias results. Hopefully, all these usability issues of VR will be solved in a near future, and VR represents a very promising tool to elicit mental workload in realistic settings like piloting or car driving (Galante, 2018) without using costly simulators. According to Santos et al. (2009), global user performances are generally better for desktop setup, partly because participants are much more familiar with this environment. However, in their experiment, participants with more computer gaming experience performed better with the VR setup. The authors argued that increasing the familiarity with the VR environment over a longer period of time will eventually allow enhanced performances.

6 CONCLUSION

Our results suggest that a virtual version of the MATB-II could provide an interesting alternative to the original 2D version. It could offer a more ecological and immersive environment for experiments in aerospace research while engaging the same type of cognitive abilities as the original version. In general, this environment was more challenging, and in some cases elicited a higher heart rate. The performance of the MATB-II VR was also very well predicted with mental rotation abilities, which can be related to the 3D environment characteristics. Some usability problems still exist in the VR environment, in particular when interacting with the small buttons of the MATB-II subtasks, but we can speculate that this issue will evolve with the improvement of the technology.

ACKNOWLEDGMENTS

This work was supported by a chair from Labsoft (Toulouse, France).

REFERENCES

Ansado, J., Chasen, C., Bouchard, S., & Northoff, G. (2021). How brain imaging provides predictive biomarkers for therapeutic success in the context of virtual reality cognitive training. Neuroscience & Biobehavioral Reviews, 120, 583-594.

Camden, A., Nickels, M., Fendley, M., & Phillips, C. A. (2017). A case for information theory-based modelling of human multitasking performance. Theoretical issues in ergonomics science, 18(3), 266-278.

Chancey, E. T., Bliss, J. P., Liechty, M., & Proaps, A. B. (2015, September). False alarms vs. misses: Subjective trust as a mediator between reliability and alarm reaction measures. In Proceedings of the Human Factors and Ergonomics Society Annual Meeting (Vol. 59, No. 1, pp. 647-651). Sage CA: Los Angeles, CA: SAGE Publications.

Chandra, S., Sharma, G., Verma, K. L., Mittal, A., & Jha, D. (2015). EEG based cognitive workload classification during NASA MATB-II multitasking. International Journal of Cognitive Research in Science, Engineering and Education, 3(1).

Daviaux, Y., Bey, C., Arsac, L., Morellec, O., & Lini, S. (2019). Feedback on the use of MATB-II task for modeling of cognitive control levels through psychophysiological biosignals. In 20th International Symposium on Aviation Psychology (p. 205).

Falah, J., Khan, S., Alfalah, T., Alfalah, S. F., Chan, W., Harrison, D. K., & Charissis, V. (2014, August). Virtual Reality medical training system for anatomy education. In 2014 Science and information conference (pp. 752-758). IEEE.

Galante, F., Bracco, F., Chiorri, C., Pariota, L., Biggero, L., & Bifulco, G. N. (2018). Validity of mental workload measures in a driving simulation environment. Journal of Advanced Transportation, 2018.

Geszten, D., Komlódi, A., Hercegfi, K., Hámornik, B., Young, A., Köles, M., & Lutters, W. G. (2018). A content-analysis approach for exploring usability problems in a collaborative virtual environment.

Hassandra, M., Galanis, E., Hatzigeorgiadis, A., Goudas, M., Mouzakidis, C., Karathanasi, E. M., & Theodorakis, Y. (2021). A virtual reality app for physical and cognitive training of older people with mild cognitive impairment: mixed methods feasibility study. JMIR serious games, 9(1), e24170.

Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., & Koo, B. H. (2018). Stress and heart rate variability: a meta-analysis and review of the literature. Psychiatry investigation, 15(3), 235.

Kennedy, L., & Parker, S. H. (2017, June). Making MATB-II medical: pilot testing results to determine a novel lab-based, stress-inducing task. In Proceedings of the international symposium on human factors and ergonomics in health care (Vol. 6, No. 1, pp. 201-208). Sage CA: Los Angeles, CA: SAGE Publications.

Krüger, J. K., & Suchan, B. (2016). You should be the specialist! Weak mental rotation performance in aviation security screeners–Reduced performance level in aviation security with no gender effect. Frontiers in Psychology, 7, 333.

Labedan, P., Darodes-De-Tailly, N., Dehais, F., & Peysakhovich, V. (2021). Virtual Reality for Pilot Training: Study of Cardiac Activity. In VISIGRAPP (2: HUCAPP) (pp. 81-88).

Lele, A. (2013). Virtual reality and its military utility. Journal of Ambient Intelligence and Humanized Computing, 4(1), 17-26.

Letondal, C., Vinot, J. L., Pauchet, S., Boussiron, C., Rey, S., Becquet, V., & Lavenir, C. (2018, March). Being in the sky: Framing tangible and embodied interaction for future airliner cockpits. In Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction (pp. 656-666).

Liu, S., & Nam, C. S. (2018). Quantitative modeling of user performance in multitasking environments. Computers in Human Behavior, 84, 130-140, p. 83.

Luong, T., Martin, N., Raison, A., Argelaguet, F., Diverrez, J. M., & Lécuyer, A. (2020, November). Towards Real-Time Recognition of Users Mental Workload Using Integrated Physiological Sensors Into a VR HMD. In 2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR) (pp. 425-437). IEEE.

Maneuvrier, A., Decker, L. M., Ceyte, H., Fleury, P., & Renaud, P. (2020). Presence Promotes Performance on a Virtual Spatial Cognition Task: Impact of Human Factors on Virtual Reality Assessment. Frontiers Virtual Real., 1, 571713.

Oberhauser, M., Dreyer, D., Braunstingl, R., & Koglbauer, I. (2018). What's Real About Virtual Reality Flight Simulation?. Aviation Psychology and Applied Human Factors.

Oberhauser, M., Dreyer, D., Mamessier, S., Convard, T., Bandow, D., & Hillebrand, A. (2015, August). Bridging the gap between desktop research and full flight simulators for human factors research. In International Conference on Engineering Psychology and Cognitive Ergonomics (pp. 460-471). Springer, Cham.

Pallavicini, F., Pepe, A., & Minissi, M. E. (2019). Gaming in virtual reality: What changes in terms of usability, emotional response and sense of presence compared to non-immersive video games?. Simulation & Gaming, 50(2), 136-159.

Pausch, R., Proffitt, D., & Williams, G. (1997, August). Quantifying immersion in virtual reality. In Proceedings of the 24th annual conference on Computer graphics and interactive techniques (pp. 13-18).

Peysakhovich, V., Monnier, L., Gornet, M., & Juaneda, S. (2020). Virtual reality vs. real-life training to learn checklists for light aircraft. In Eye-Tracking in Aviation. Proceedings of the 1st International Workshop (ETAVI 2020) (pp. 47-53). ISAE-SUPAERO, Université de Toulouse; Institute of Cartography and Geoinformation (IKG), ETH Zurich.

Santiago-Espada, Y., Myer, R. R., Latorella, K. A., & Comstock Jr, J. R. (2011). The multi-attribute task battery ii (matb-ii) software for human performance and workload research: A user's guide.

Santos, B. S., Dias, P., Pimentel, A., Baggerman, J. W., Ferreira, C., Silva, S., & Madeira, J. (2009). Head-mounted display versus desktop for 3D navigation in virtual reality: a user study. Multimedia tools and applications, 41(1), 161-181.

Singh, R. P., Javaid, M., Kataria, R., Tyagi, M., Haleem, A., & Suman, R. (2020). Significant applications of virtual reality for COVID-19 pandemic. Diabetes & Metabolic Syndrome: Clinical Research & Reviews, 14(4), 661-664.

Sladky, R., Stepniczka, I., Boland, E., Tik, M., Lamm, C., Hoffmann, A., & Windischberger, C. (2016). Neurobiological differences in mental rotation and instrument interpretation in airline pilots. Scientific reports, 6(1), 1-6.

Smith, V., Warty, R. R., Sursas, J. A., Payne, O., Nair, A., Krishnan, S., & Vollenhoven, B. (2020). The effectiveness of virtual reality in managing acute pain and anxiety for medical inpatients: systematic review. Journal of medical Internet research, 22(11), e17980.

Soret, R., Charras, P., Hurter, C., & Peysakhovich, V. (2019). Attentional orienting in virtual reality using endogenous and exogenous cues in auditory and visual modalities. In Proceedings of the 11th ACM Symposium on Eye Tracking Research & Applications (pp. 1-8).

Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. Behavior

Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. Teaching of Psychology, 44(1), 24-31.

Symonenko, S., Zaitseva, N., Osadchyi, V., Osadcha, K., & Shmeltser, E. (2020). Virtual reality in foreign language training at higher educational institutions.

Wismer, P., Cordoba, A. L., Baceviciute, S., Clauson-Kaas, F., & Sommer, M. O. A. (2021). Immersive virtual reality as a competitive training strategy for the biopharma industry. Nature Biotechnology, 39(1), 116-119.

Zygouris, S., Ntovas, K., Giakoumis, D., Votis, K., Doumpoulakis, S., Segkouli, S., & Tsolaki, M. (2017). A preliminary study on the feasibility of using a virtual reality cognitive training application for remote detection of mild cognitive impairment. Journal of Alzheimer's Disease, 56(2), 619-627.