The Superiority Effect of Virtual Reality on Prospective Memory

by

Jinous Mirzaagha

A thesis submitted to the Faculty of Graduate and Postdoctoral Affairs in partial fulfillment of the requirements for the degree of

Masters of Applied Science

in

Human-Computer Interaction

Carleton University
Ottawa, Ontario

Abstract

Although failures of prospective memory have been linked to aviation accidents, prospective memory has not been integrated into pilot training or assessment. Virtual Reality (VR) offers ecologically robust flight environments for training and assessment of prospective memory, but the effects of VR, as compared to 2D displays, have yet to be studied. VR flight simulation is also an ideal environment for studying the detection of time- or event-based cues, which are key to triggering the prospective memory task. Participants (N=25, non-pilots) completed both time- and event-based prospective memory tasks along an automated route, with the percentage of completed tasks recorded. Participants were trained on the prospective memory cues in either VR or 2D but "flew" in both conditions: VR (head-mounted 3D device), and 2D (flat screen display). Results from repeated measures analysis of variance tests showed that performance was enhanced in VR only for time-based prospective memory tasks.

Acknowledgements

Thank you to Dr. Chris Herdman for the opportunity to join the Advanced Cognitive Engineering Lab. To Dr. Kathleen Van Benthem for her support and guidance throughout this process. To James Howell for contributing his technical expertise. Finally, to all of my wonderful friends and colleagues in the ACE lab for making this experience so fun.

Table of Contents

List of Tables

List of Figures

List of Appendices

This page lists all of the appendices.

Introduction

Flight simulators are an essential tool for training and assessments for pilots. Flight simulators are safer than flying real aircraft and provide researchers with control over the environmental factors that are being measured. VR flight simulators in particular, provide many advantages over other types of flight simulators, such as full-scale replications, because they are typically less expensive, more accessible, and portable. Therefore, using a VR system may provide solutions to some of the practical issues commonly encountered in aviation research.

There is recent evidence that the ecological validity of VR environments makes them ideal testbeds for cognition research (Kourtesis et al., 2020). However, further research is needed to understand how effective technology is for training and assessment, and whether VR could replace traditional flight simulators for strategies designed to reduce or manage pilot cognitive errors.

One cognitive factor that is linked to pilot error is prospective memory, that is remembering to perform a task in the future. Not only does prospective memory support aviation, navigation, and communication during flight, prospective memory performance also predicts the likelihood of a pilot incurring a critical incident (Van Benthem & Herdman, 2020). Prospective memory is a complex cognitive function where performance is tied to the detection and type of triggers, or cues (a perceptual trigger to perform an intended action), in the environment. Whether a cue is event-based or time-based (see Figure 1) can affect its detection. Since most prospective memory intentions require that cognitive functions for monitoring the environment or internal time-keeping overlap the ongoing activities associated with flying an aircraft, prospective memory performance is susceptible to moderation by individual factors, such as age and experience.

The activation model of prospective memory is described in terms of how the memory is accessed under different conditions. A memory item may be activated based on the distinctiveness and salience of the cue, and an item that has greater activation can therefore be accessed (Dismukes, 2010; Brandimonte & Passolunghi, 1994). An example of distinctiveness is provided by Dismukes (2010) in an experiment using target words, where an unusual word (unusual in both semantic and visual appearance) yields greater activation power than a familiar word, because they are a more novel cue that may alert the individual that has been habituated to similar cues. Additionally, Dismukes 2010 describes that fewer prior associations to the unusual word, allow the cue to act in a more direct and targeted way towards the desired prospective memory intention. In identifying factors that influence prospective memory, there have been a number of suggested ways to design human-machine interfaces to mitigate the sometimes catastrophic results of prospective memory failures.

The present study aims to understand if prospective memory activity can be identified through the ecologically valid method that VR flight simulators afford, and if so, future studies may be able to incorporate in the VR interface a number of methods that have been successful in mitigating prospective memory failures. Dismukes (2010), lists a number of potential methods for both the designer and the individual. For the designer, it is suggested that careful consideration is taken to analyze environmental “hot spots”, where prospective memory failures are likely to occur, and thereby anticipate failure, and implement safety protocols.

Substantial prior research comparing 2D and 3D displays show VR superiority for many different tasks. In their review of almost 50 years of research, McIntire, Havig and Geiselman (2014) found support for the premise that 3D displays result in better performance, particularly in object-manipulation and complex and novel tasks when compared to 2D displays. McIntire et al. reported reduced or similar performance in VR conditions when the tasks were simple or when navigation was required.

Given that previous research has demonstrated that increase workload demands may have an inverse effect on prospective memory (Van Benthem, 2015), one of the suggested methods of preventing prospective memory may be to reduce demands by implementing strategically organized data communication, checklists, warning systems, and alerting systems to keep track of tasks. Dismukes (2010), also suggests that following further research in prospective memory, educational initiatives may assist in making the individual aware when prospective memory failure are likely to occur and how to protect themselves. The methods for individuals include creating self-reminder cues, avoiding deferral of tasks, minimize multitasking where possible, particularly if a vital task is included. Additionally, given the previous discussion on the importance of habitual tasks in automaticity, it may be beneficial to link non-habitual tasks to these tasks so that they are attached.

Prospective Memory

Prospective Memory is an emerging field of study and is defined as the ability to hold in memory an intention to carry out a task in the future. Prospective memory can be triggered by external factors known as event-based cues, for example remembering to do something because you see something in the environment that is associated with the task, such as remembering to make a radio call when a certain altitude is reached. Prospective memory can also be triggered by internally tracked factors such as time and are known as time-based prospective memory cues, for example remembering to check the oil pressure every five minutes.

Prospective memory and retrospective memory can be difficult to differentiate, and there is some overlap in their presentation. Retrospective memory differs from prospective memory because retrospective memory relies on recall. However, prospective memory also requires retrospection, in that the participant must remember the intention. In prospective memory, the intention to remember must occur prior to the recall, and in retrospective memory, the intention to remember must occur at the time of recall. In aviation for example, pilots are required to carry out complex procedural tasks and often face interruptions during a flight. Limitations in cognitive resources require the pilot to perform procedural tasks at the right time. However, prospective memory requires a different methodology to evaluate. With prospective memory in the real world, there may not be an explicit cue to trigger the prospective memory, as the individual may not be actively trying to remember, as someone might be when they are attempting to retrieve a stored memory intentionally. The difference in laboratory and naturalistic settings has been partially explained by Marsh et. al (2005), who suggest that there is an interaction between closely and non-closely associated cues to their response, and whether they are actively monitoring the cues during the response phase. Therefore, it is important to research prospective memory in terms of the cognitive attributes associated with both successful and unsuccessful prospective memory retrieval, and to find what are the cognitive and environmental factors that make them successful or not. Stages of prospective memory are described by Dismukes 2010 as: 1) an encoding stage where the intention to perform a task at a later time is formed, 2) a retention stage where continual awareness of the intention may be interrupted by other ongoing tasks, and 3) a retrieval phase where the intention is retrieved, and the action takes place.

Time- and Event- based Prospective Memory

Time-based prospective memory requires remembering to complete a planned action after a certain time has passed, for example remembering to complete a task in a few hours. Event-based prospective memory requires remembering to complete a planned task after some event has occurred, for example encountering an associated landmark.

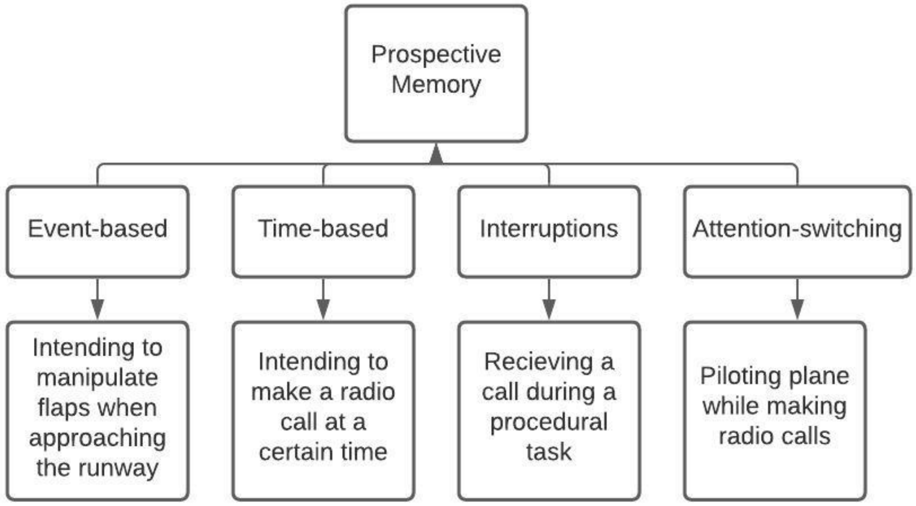


Figure 1: Types of prospective memory.

Prospective Memory in Aviation

Research into prospective memory is important in aviation, as pilots are frequently called upon to delay intentions, and carry them out in the future amid high workload ongoing contexts. Prospective memory plays an important role in keeping track of complex aviation tasks required during flight. An overview of the causes of aviation accidents found that many accidents were caused by experienced pilots missing a procedural step that they had successfully executed before (Dismukes, 2007).

Previous research has suggested that there may be countermeasures to potentially reduce the impacts of prospective memory failure in pilots (Dismukes, 2010). Aviation specific tasks are particularly applicable in investigating prospective memory because pilots are required to monitor many different instruments, and communications simultaneously. Some of the tasks that pilots regularly do become habitual and automatic, however other tasks that occur less frequently as interruptions do not have the opportunity to become habitual, and therefore task switching onto and away from the one-off occurrences become susceptible to increased prospective memory failure. Pilots may also become engaged in monitoring one area, such as the instrument panel, radio communications, or field of view through the window when flying low to the ground, and during these high engagement periods, pilots are vulnerable to what is described by Dismukes (2010) as cognitive tunneling.

There are a number of ways that flight procedures may be prone to errors. One of the ways is through pilot error in monitoring the environment sufficiently to receive information about the tasks they should do. Another is that the cue is not recognized as a cue and therefore does not act as a trigger. Third, given the multi-step process required for the procedural nature of the task list, there may not be a proper formation of the intent to perform the next necessary steps. For example, Van Benthem (2015), investigated how factors such as (memory cue salience and workload) and individual differences (pilot age, cognitive health, and expertise) impacted prospective memory for communication tasks in general aviation. There were 101 trained pilots flying a circuit and completing memory tasks while maintaining communication. The study found that prospective memory was negatively impacted on tasks that used low salience cues in high workload conditions. The results suggest that failures in prospective memory may be mitigated through technologies that can help increase cue salience or assist in memory during anticipated high-workload conditions.

There have been multiple ways to mitigate these issues, including using flight aids, checklists, however, flight aids may require greater attentional resources to be taken away from monitoring other environmental features, and unexpected interrupts may lead to critical errors in executing check lists effectively.

Virtual Reality and Prospective Memory

VR is a technology that may facilitate the study of prospective memory in an ecologically valid way. VR is used in aviation research because of its many advantages over full-scale simulations, or real aircrafts. Using full-scale simulations and aircrafts can be costly and inaccessible to many. However, VR is relatively accessible, and advances in VR technology allows the ability to achieve testing without requiring large spaces or too much equipment. Furthermore, VR can be a safer alternative to using real-world scenarios, as participants are not exposed to potential dangers such as crashes, or environmental factors that may lead to unsafe flying conditions, such as low visibility. Finally, VR provides the opportunity to manipulate the prospective memory tasks in order to have a greater level of control over the output measures in studying PM.

One of the reasons that prospective memory may afford this opportunity is the importance of distinctive visual cues in event-based prospective memory. In time-based prospective memory, there is a lower reliance on visual distinctiveness, and a greater reliance on familiarity, and research on how VR may enhance time-based prospective memory is less supported than in event-based prospective memory.

Decades of research into the benefits of 3D and VR over 2D displays, McIntire, Havig & Geiselman (2014) indicate that prospective memory performance may also experience the VR superiority effect. Due to the congruence of the focality (locus of attention) of the event-based cues with the focus of the ongoing aviator tasks, it was expected that event-based prospective memory would exceed performance for time-based PM.

Additionally, Wagner et. al. (2019) utilized a Space-Time Cube prototype environment where the cube is coupled to a 3D virtual representation of a desk, and space and time are intuitively controlled through gestures. The prototype was compared to a 2D desktop environment (standard mouse and keyboard controls). Their findings support a VR superiority effect as demonstrated by some advantages in their 3D interface in comparison to their 2D interface. The research hypothesized that the immersive prototype would lead to faster task completion time, greater accuracy, greater intuitiveness, lower mental workload, and higher usability ratings, and found some evidence to confirm differences in 3D over 2D.

Previous research has shown that when a variety of tasks are completed in 2D and VR environments, there is some evidence that VR environments afford an enhanced performance score. The finding that VR provides enhanced prospective memory performance over 2D is supported by studies that look at immersive 3D interfaces and their improved usability when compared to standard 2D interfaces.
When the stimuli linked to prospective memory performance have multidimensional features, the advantages of VR may be enhanced. For example, research on radar signal

detection found that 3D immersion allowed for multidimensional information to be analyzed better, when compared to 2D signals, because the information is presented in a way that is cohesive rather than presented as separate entities (Cantu et al., 2019). This study demonstrates the increased interest in utilizing VR superiority and technology to enhance visual tasks, although the task is not relevant to a typical general aviation task, the results do support that there may be some benefit in utilizing a 3D interface in order to manipulate cue presentation, that has previously been demonstrated (Dismukes 2010; Einstein & McDaniel 2004, Van Benthem, 2015) as a vital component in investigating how cue distinction and display may impact its saliency power.

In previous research, Stefano et. al. (2017) were primarily concerned with the first or third-person perspective of how the map is displayed and compared both perspectives through a user study involving wayfinding tasks. The motivation of this study comes from the usefulness of mobile 3D maps to facilitate navigation in large indoor spaces. The perspectives include: a mobile 3-D map with first-person perspective, a mobile 3-D map with third-person perspective, and a traditional mobile 2-D map. The research found participants were able to reach their destinations in all three conditions without errors the idea that the 3rd person perspective allows users to match a 3D map with the physical world. Overall, third-person perspective led to shorter orientation time before walking, better clarity ratings, lower workload, mental demand and effort scores, and higher preference score compared to the mobile 3-D map with first-person perspective.

Additionally, it has been found that immersive environments demonstrate an ability to increase presence in virtual environments, and thereby increase physiological response from users in the environment (Etemadpour, Monsoon & Linesen, 2013). This work investigated whether a projection onto a 3D physical space could improve performance when compared to the projection on a 2D physical space. They hypothesized that when they display a projection into a 3D visual space it can increase the performance of common visual analysis tasks due to a higher projection precision, when compared to the projection of the 3D visual space on a 2D screen. In the Eetemadpour et al. study participants were presented with different fidelity conditions (HD or LD), they were asked to view a projection of a cluster of images and to complete tasks such as count the clusters, detect specific clusters, and identify density. They confirmed that distances between individual objects can be perceived better in VR that led to better overall performance in local tasks only (where they only looked at specific parts of the visual, not global where they require comprehending the distribution of the entire scatterplot). When measuring whether HD can outperform LD, LD never showed significantly higher accuracy than HD, however HD only had a few tasks where it had significantly higher accuracy, they still concluded that HD could improve performance. They also found that users had a slight preference for VR.

Environmental Specificity Principle

Episodic memory has been defined as an alternative to semantic memory (Tulving, 1972). Semantic memory is memory that is generated through general experience of the world, episodic memory is a personal form of memory association.

An example of an episodic memory has been found in studies that demonstrate that participants perform better on an exam if the exam was taken in the same environment that the participant studied for it in. In a study with professional divers, (Godden & Baddeley, 1975), divers were given a list of words to memorize in two environments: on dry land and underwater. The divers that studied underwater performed better at recalling the words when underwater, and the divers that learned on dry land performed better at recalling the words when on dry land.

Subsequent studies sought to understand if the results found in divers could be generalized to other contexts. In medical education, participants were asked to learn a list of words in either a clinical (bedside) environment or an educational (classroom) environment and found that aspects of the learning task were performed better in congruent learning-to-test environment (Koens, Ten Cate & Custers, 2003).

Hypothesis

Virtual Reality Superiority

It is expected that prospective memory performance in the 3D environment would be superior to prospective memory performance in the 2D environment. This hypothesis is based on literature showing that VR provides features that increase immersion and may positively affect PM. Given that in 3D VR immersion is enhanced via higher fidelity and participant engagement, it was expected that participants would perform better overall in the 3D VR environment.

It is possible that performance could be enhanced in tasks where training occurred in the same environment that participants were also tested in. This potential moderator of VR superiority is supported by the environmental specificity principle (Tulving, 1973).

Cue Type

It is expected that performance would be better for prospective memory tasks that used event-based cues rather than time-based cues. In an aviation environment it would be expected that performance in event-based prospective memory tasks would be impacted by the environmental differences, due to the visual nature of event-based cues, in comparison to time-based cues that rely more on internal time-keeping, this hypothesis is supported through previous research in aviation that demonstrates the importance of visual cues, and their priority over internal time-keeping. The interaction between Test Graphics (VR vs. 2D) and Cue Type will also be investigated to determine whether the VR condition enhances performance for the time-based or the event-based prospective memory task. If visual cues are more pronounced in VR then the interaction may show enhanced performance in the event-based condition. However, if situation or time awareness is enhanced, then the time-based and event-based prospective memory performance may be similar in the VR condition.

Method

Participants

There were 25 total participants from the undergraduate research pool. The participants were not required to have pilot training, and were only asked to have normal to corrected vision and hearing.

Design

The analysis looked at the Test Graphics (2D or 3D test conditions), the Training Environment (2D or 3D test conditions). A 2 (first factor) 2x2x2 repeated measures ANOVA test will be conducted.

Materials

There were two environments where the participants were placed. A standard 2D flight replica environment (see Figure 6) that included: 3 computer monitors, a yoke, a physical number pad, and a flap. A 3D environment, that included an HTC Vive headset, a virtual number pad, a virtual flap (that is spatially located above the physical flap). In both environments, participants heard tones playing for an auditory peripheral detection test but were not required to wear any headsets to hear the tone.

Procedure

Participants were shown two legs of a pre-recorded flight where they were asked to perform both time-based and event-based prospective memory tasks throughout the task, following a practice session in either a standard two-dimensional model, or a three-dimensional VR mode (see Figure 5). The recording of the flight scenario was beneficial because it allowed for equality among participants in their ability to fly. The participants were focused on the prospective memory tasks, which controlled when the participants were exposed to certain cues. Participants were given an initial functional training phase to make them acquainted with the hardware. All participants were then asked to complete the two automated test routes (aircraft controls on auto-pilot), in both 2D and VR. The test flights were counterbalanced, the first and second leg of the flight were counterbalanced to be encountered in both VR and 2D environments to account for learning effects (see Figure 2).

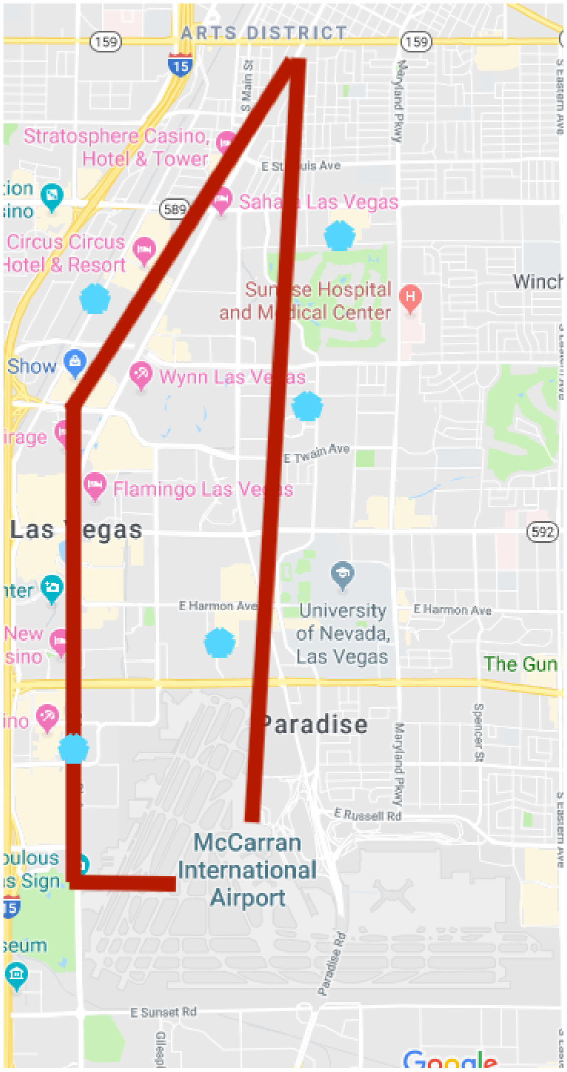


Figure 2: The flight route taken in each Test Condition. Participants flew either clockwise or counter-clockwise in the first session and the opposite direction in the second session for counterbalance.

The study compared the two display modes of flight simulation interfaces; a 3-monitor non-VR interface with a physical keyboard, and a VR interface with the HTC Vive headset, combined with virtual keypads. Both simulation modes incorporated flaps, yoke, and seat replicated from a Cessna aircraft.

In the VR set-up, participants wore the headset and did not make use of the monitors, they were given a virtual keypad through their viewing lens and could look around the environment with a 360° view, the instrument panel in VR was set within the cockpit.

In the non-VR set-up, participants were given three standard display monitors, and were provided a physical keypad. Participants were able to look around their environment but were asked to rotate the yoke to change the view. The instruments in the non-virtual set-up were placed on the screen above the cockpit panel, to limit the effort in changing view and to reduce the advance that participants in VR had in viewing by gaze.

Following a briefing, participants were provided with instructions on the flight tasks and the functions of the components of the flight simulator. Participants were guided through three practice modules. In the SONA student group, half of the participants were given the practice sessions in standard two-dimension, the other half, will be given the same practice session in three-dimension VR mode. The practice modes will be set in Ottawa, Ontario.

In the first practice module, the participants practiced their landmark recognition task, and will be required to remember the Call Sign given by nearby aircraft and type the Call Sign when they approach the Bellagio landmark.

In the second practice module the participants toggle the flaps after each minute has passed.

In the third practice module the participants entered the airspeed indicated in the instruments, when a sudden altitude change occurs, through both features of the environment, and changes on the altimeter. Participants repeated the practice module if there are any errors or delays in completing the practice task.

After completing the practice modules, the participants were asked to complete an approximately 10-minute flight. Each participant was tested in the standard two-dimensional flight mode, then given a route that traveled in reverse, with the same tasks and features in the three-dimensional VR mode.

The tasks included an event-based prospective memory task. The event-based prospective memory task required the participants to type into the keypad, a previously heard Call Sign once they encounter the event. The event was marked by a landmark fountain (see Figure 3) that the participant hovered over. Completing the event-based prospective memory task required that the participant monitor the environment for an aircraft that enters the airspace, and monitor the auditory information provided by the Call Sign. They were also required to monitor the environmental features to identify the Bellagio.

There was a time-based prospective memory task that required the participant to toggle the flap after every minute that passes. Completing the time-based prospective memory task, required the participant to monitor the timer, as they continue to monitor other task features and instruments. They were given additional notice to time after every second minute has passed, that will provide a more salient cue.

Participants were asked to remain aware of their altitude. During the flight they will experience a sudden change in altitude, and participants were asked to type in their airspeed during this time. The change in altitude will provide a lower salient cue that is inherent with sudden altitude changes, where the environment and speed will change. The higher salient cue will be provided by the altimeter instrument, that they were asked to monitor and respond when it reached the required altitude described to them during the briefing.

There was a peripheral detection test during the flight, where participants were asked to respond to a tone by pressing the yoke. The auditory PDT was a tone played peripherally throughout the flight in order to confirm that the 2D and VR environments did not differ significantly in difficulty and workload (see Appendix E). A 1000hz tone was played at a comfortable volume and occurred in random intervals of 4 to 7 seconds. Participants were asked to respond to the tone by pressing a button on the yoke. The participants were scored as successful if there was a button press following the presentation of a tone.

Figure 3: The fountain as a cue for participants to enter the call-sign previously heard.

Dependent Variables

Time-based prospective memory task (Flaps)

Pilots were provided with a timer in the cockpit (see Figure 4). Participants were asked to monitor the timer as they completed other tasks simultaneously, and to toggle the physical flap after they noticed a minute had passed on the timer. Participants were also given a visual reminder after every second minute had passed to toggle the flap. The visual reminder was incorporated as an event-based cue to measure any potential difference between time-based and event-based cues. The visual reminder would appear on the screen in focal view in both 2D and VR environments. Participants were scored if in each minute the flap was toggled to the alternate position.

For the time-based task, participants were scored out of 10 possible completions, and a point of completion was awarded for every flap-down motion that was tracked through a data collection automated program.

Figure 4: (Top) An image of the instrument panel. (Bottom) An illustration of the flap as displayed in the VR mode)

Combined event-based prospective memory task (Altimeter/Call-sign)

Altimeter

During each flight, the aircraft climbed to an altitude of over 3000 feet, at this point the participants were cued to the altitude change by changes in the environment (zero ground visibility), and the altimeter instrument that they were asked to monitor. Participants were told during the training period that when the given altitude is reached they will look at their airspeed in the instrument panel and enter that speed into the keypad. In the 3D environment, they would enter the airspeed into the virtual keypad, and in the 2D environment, they would enter it into a physical keypad. Participants were successful if they entered the speed during the change in altitude.

For the Altimeter task participants received one point for each digit pertaining to their speed that they input within 15 seconds of crossing the altitude. Scores for this task ranged from a potential of 0 points to 2 points based on the accuracy of the speed entered. The two event-based tasks were combined for greater variability in scores.

Call-sign

An aircraft would pass through the airspace and provide a Call Sign, when the fountain would appear in view, participants were asked to enter the Call Sign number when the fountain was visible to them. The aircraft would hover over the fountain (see Figure 3).

For the Call-sign task, participants received a point for each correct number they input into the keypad – one for each of the aircrafts they encountered (2). Therefore, there was a possible minimum score of 0 points and a maximum score of 6 points. A list of all tasks can be found in Table 1.

Independent Variables

Test Graphics

The Test Graphics were the 2D and 3D modes, to investigate the effect of the Training Environment and the Test Graphics on prospective memory performance, a repeated measures ANOVA was conducted with one within subject factor (Test Graphics) and one between subject factor (Training Environment).

Training Environment

The Training Environments were the 2D and 3D modes that the participants were first trained in using two within subject factors (Test Graphic and Cue Type) and one between subject factors (Training Environment).

Table 1

Prospective Memory Task List

Task 1	Flaps	Time-Based	Participants respond to time throughout the flight.
Task 2	Call-Sign	Event-Based	Participants respond to an auditory Call-Sign.
Task 3	Altimeter	Event-Based	Participants respond to altitude by instrument panel and zero ground visibility.

Table 1: Prospective Memory Tasks

Figure 5: Virtual view while wearing HTC Vive headset.

Figure 6: The simulated cockpit with yoke, flap, and toggle, as found in a typical Cessna aircraft.

Results

Test Graphics and Training Environment

To investigate the effect of Test Graphics on prospective memory, a repeated measures ANOVA was conducted with one within-subject factor (Test Graphics: 2D or VR) and one between-subject factor (Training Environment: 2D or VR). Outcomes were measured in terms of the percentage of correctly completed prospective memory tasks. This analysis was done as an exploratory analysis, thus the significance threshold was set at $p < .1$ with an effect size >.1 (Eta-squared [$\eta 2$] or partial eta-squared [$\eta p2$]). All analyses met the assumptions of repeated measures ANOVAs, including independent observations and multivariate normal distribution in the population. The assumption of sphericity (that the variances of all difference scores among the test variables must be equal in the population) does not apply in the following analyses as outcome variables were repeated only once.

Time-based Prospective Memory

As shown in Figure 8, there was a significant main effect of Test Graphics for the time-based prospective memory variable such that in the 2D condition ($M = 86$) there was lower performance than in the VR condition ($M = 94$), $F(1, 23) = 6.61$, $p = 0.017$, $\eta p2 = 0.223$. No main effect of Training Environment was found for the time-based Flaps prospective memory variable, $p = 0.356$, $\eta p2 = 0.037$. Although Figure 7 appears to show a trend, the interaction between Test Graphics and Training Environment on the time-based prospective memory variable was not significant $F(1, 23) = 0.886$, $p = 0.629$, $\eta p2 = 0.010$). VR superiority was seen with time-based prospective memory and this effect was not moderated by the Training Environment. For this time-based prospective memory task, support for the VR superiority hypothesis was shown by finding that the lowest performance was in the 2D condition.

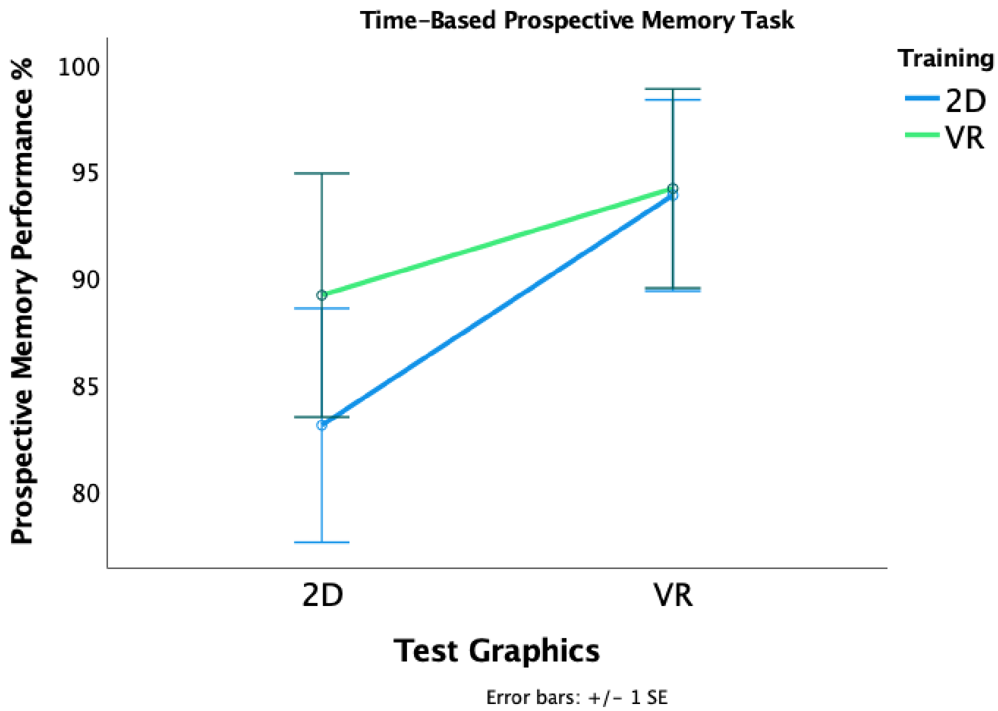


Figure 7: Main effect of graphics type for the time-based Flaps prospective memory variable.

Event-based Prospective Memory

No main effect of Test Graphics was found for the event-based Airspeed/Call Signs prospective memory task, As shown in Figure 8, the mean 2D (M = 90.50) and VR (M = 90) percentage of completed tasks were virtually equal across graphics conditions. There was no main effect of Training Environment, $p = 0.853$ and no significant interaction between Test Graphics and Training Environment, $p = 0.440$.

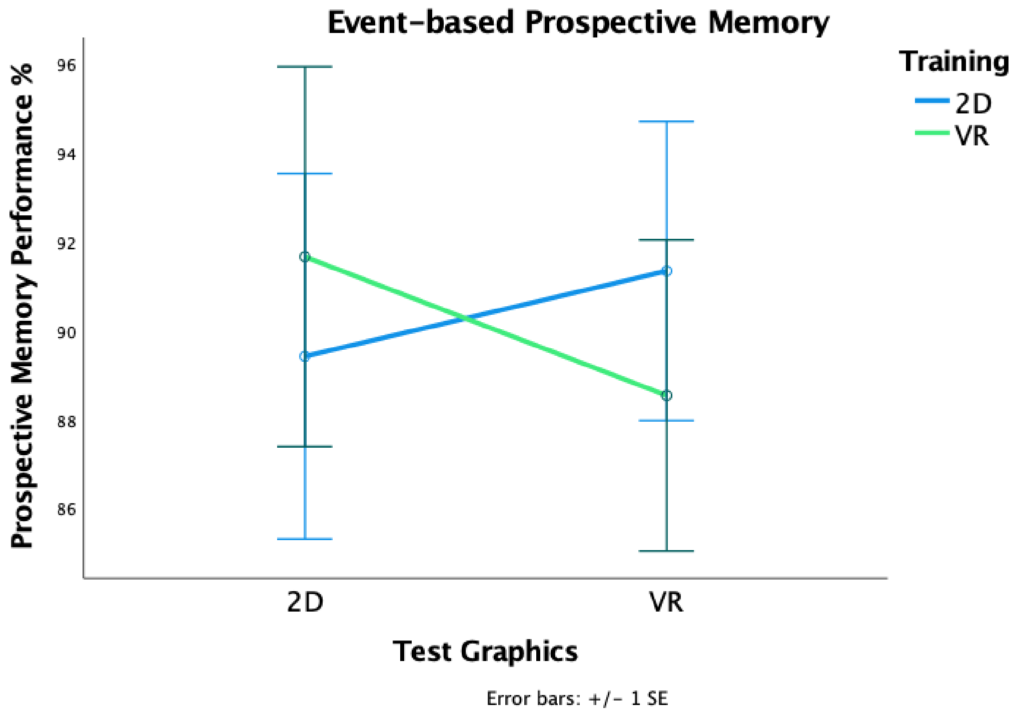


Figure 8: No main effect of Test Graphics for the event-based Altimeter/Call Sign prospective memory.

Cue Type

Time-based and event-based cues have been shown to elicit differences in prospective memory performance. In this design, a two-factor repeated measures ANOVA was conducted to quantify the effect of Cue Type (time- vs. event-based) on prospective memory and to investigate whether any effects of Cue Type might interact with Test Graphics (2D vs. VR). Outcomes were measured in terms of the percentage of completed prospective memory tasks. As this test was also exploratory, the significance threshold was set at $p< .1$ with an effect size >.1 (Eta-squared [$\eta 2$] or partial eta-squared [$\eta p2$]). All analyses met the assumptions of repeated measures ANOVAs, including independent observations and multivariate normal distribution in the population. The assumption of sphericity (that the variances of all difference scores among the test variables must be equal in the population) does not apply in the following analyses as outcome variables were repeated only once.

As shown in Figure 9, there was an expected main effect of Test Graphics $F(1, 23) = 4.024$, $p = 0.05$, $\eta p2 = 0.144$. There was no main effect of Cue Type 2D-event-based ($M = 86$), 3D Event-Based ($M = 94$), 2D Time-Based ($M = 90.5$), 3D-TimeBased ($M = 90$). Also shown in Figure 9, there was a marginal interaction between Test Graphics and Cue Type found, $F(1, 23) = 1506$, $p = 0.00$, $\eta p2 = 0.984$. This interaction between Test Graphics and Cue Type, where only the time-based prospective memory performance was enhanced via the VR graphics type (see Figure 9). The interaction was further analyzed using a Bonferroni post-hoc analysis, which showed that the difference in time-based performance across the two graphics conditions was significant, ($p = 0.015$).

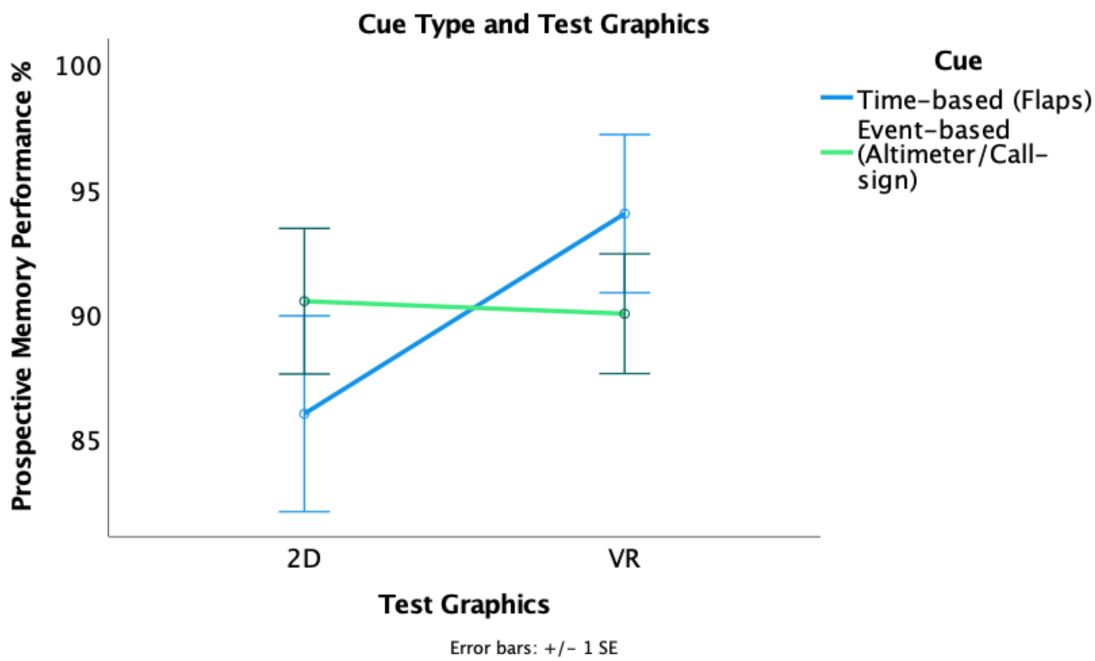


Figure 9: Interaction between Cue Type and Test Graphics.

Discussion

VR Superiority

The results of the present research have shown that VR may impact prospective memory under some conditions. Although we had hypothesized that VR superiority would produce enhanced performance in event-based cues because of the visual nature of event-based cues, and the visual differences between 2D and VR environments. However, the results show that it was time-based prospective memory tasks that were enhanced in a VR environment. There are two potential explanations for this finding

The first possible explanation is that the VR environment may enhance perceptual triggers, and participants in a VR environment may free up mental resources in monitoring the environment, and therefore have greater mental resources to focus on time-keeping. This explanation is supported through previous work that demonstrates that workload may impact situational awareness and participants in high workload conditions, have been shown to have worse prospective memory performance (Van Benthem, 2015). However, the present research conducted an auditory peripheral detection task and found that the workload in both 2D and VR were not significantly different. Therefore, an alternative explanation may be more appropriate.

The second possible explanation may be that improved prospective memory performance in time-based tasks is derived from factors that may be enhanced in VR such as immersion and situational awareness, these aspects could lead to a greater awareness of time. This explanation is supported through previous research that has found a strong correlation between low levels of situational awareness, and performance measured by the number of errors produced when participants controlled the system (Endsley and Kiris, 1995).

Given the robust findings that show correlations between situational awareness and performance in complex dynamic tasks in aviation, research in situation awareness has more recently been conducted in non-aviation research as well (Gaba & Howard, 1995), that found that situational awareness can impact specific aspects such as picking up on subtle cues, situations that evolve and are unpredictable, and elements that require special knowledge. Our results demonstrate that having a greater awareness that is afforded through the VR environment, may increase sensitivity to cues, including internal time-keeping cues.

Conclusion

The aim of this research was to explore the differences in how prospective memory works in 2D and VR environments. In order to understand these features, we made two main hypotheses that predicted that prospective memory performance would be enhanced overall (fewer errors made) in a VR environment when compared to a 2D environment, and that the environment where training first occurred may impact these results. Secondly, we predicted that there would an effect of cue type, where event-based prospective memory would benefit more from the VR environment. Our results showed that there was no overall effect of VR superiority, but that time-based prospective memory was enhanced when tested in a VR environment. It is predicted that time-based advantage is provided through a potentially greater situational awareness in a more immersive environment.

In the future it would be beneficial to follow-up these results with a deeper investigation into event- vs. time-based prospective memory, and situational awareness in VR. Recently we have seen an increase in VR being utilized in many different applications, and it is therefore a worthwhile area of research to understand how prospective memory is affected in VR environments. Given that failures in prospective memory, and lower situational awareness may lead to catastrophic results in not only the aviation field, but the health and safety field as well, research in this area may help to minimize these failures in VR applications.

Bibliography

Brandimonte, M. A., & Passolunghi, M. C. (1994). The effect of cue-familiarity, cue-distinctiveness, and retention interval on prospective remembering. *The Quarterly Journal of Experimental Psychology A: Human Experimental Psychology*, *47A*(3), 565–587.

Cantu, A., Duval, T., Grisvard, O., & Coppin, G. (2019). *Expert Evaluation of the Usability of HeloVis: A 3D Immersive Helical Visualization for SIGINT Analysis*.

Dismukes, K. (2010). Remembrance of Things Future: Prospective Memory in Laboratory, Workplace, and Everyday Settings—R. Key Dismukes, 2010. *Reviews of Human Factors and Ergonomics*.

Dismukes, K., & Nowinski, J. (2007). Prospective memory, concurrent task management, and pilot error. In *Attention: From theory to practice* (pp. 225–236). Oxford University Press.

Einstein, G. O., & McDaniel, M. A. (2004). *Memory fitness: A guide for successful aging* (pp. xiv, 290). Yale University Press.

Einstein, G. O., McDaniel, M. A., Thomas, R., Mayfield, S., Shank, H., Morrisette, N., & Breneiser, J. (2005). Multiple processes in prospective memory retrieval: Factors determining monitoring versus spontaneous retrieval. *Journal of Experimental Psychology. General*, 134(3), 327–342.

Endsley, M. R., & Kiris, E. O. (1995). The Out-of-the-Loop Performance Problem and Level of Control in Automation. Human Factors, 37(2), 381–394.

Etemadpour, R., Monson, E., & Linsen, L. (2013). The Effect of Stereoscopic Immersive Environments on Projection-Based Multi-dimensional Data Visualization. *International Conference on Information Visualisation*.

Gaba, D. M., Howard, S. K., & Small, S. D. (1995). Situation awareness in anesthesiology. Human factors, 37(1), 20-31.

Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, *66*(3), 325–331.

Koens, F., Ten Cate, O., & Custers, E. (2003). Context-dependent memory in a meaningful environment for medical education: In the classroom and at the bedside. *Advances in Health Sciences Education : Theory and Practice*, *8*.

Kourtesis, P., Korre, D., Collina, S., Doumas, L. A. A., & MacPherson, S. E. (2020). Guidelines for the Development of Immersive Virtual Reality Software for Cognitive Neuroscience and Neuropsychology: The Development of Virtual Reality Everyday

Assessment Lab (VR-EAL), a Neuropsychological Test Battery in Immersive Virtual Reality. *Frontiers in Computer Science*, *1*.

Marsh, R. L., Hicks, J. L., & Cook, G. I. (2005). On the Relationship Between Effort Toward an Ongoing Task and Cue Detection in Event-Based Prospective Memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*(1), 68–75.

McIntire, J., Havig, P., & Geiselman, E. (2014). Stereoscopic 3D displays and human performance: A comprehensive review. *Displays*, *35*, 18–26.

Oates, J. M., & Peynircioğlu, Z. F. (2014). The importance of source and cue type in time-based everyday prospective memory. *The American Journal of Psychology*, *127*(3), 343–350.

Smith, R. E., & Bayen, U. J. (2005). The effects of working memory resource availability on prospective memory: A formal modeling approach. *Experimental Psychology*, *52*(4), 243–256.

Stefano, B., Chittaro, L., & Sioni, R. (2017). Mobile Three-Dimensional Maps for Wayfinding in Large and Complex Buildings: Empirical Comparison of First-Person Versus Third-Person Perspective. *IEEE Transactions on Human-Machine Systems*, *47*.

Tulving, E. (1972). Episodic and semantic memory. In *Organization of memory* (pp. xiii, 423–xiii, 423). Academic Press.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373.

Van Benthem, K. D., Herdman, C. M., Tolton, R. G., & LeFevre, J.A. (2015). Prospective memory failures in aviation: Effects of cue salience, workload, and individual differences. *Aerospace Medicine and Human Performance*, *86*(4), 366–373.

Wagner, J., Stuerzlinger, W., & Nedel, L. (2019). Evaluating an Immersive Space-Time Cube Geovisualization for Intuitive Trajectory Data Exploration. *IEEE Transactions on Visualization and Computer Graphics*.

Appendices

Appendix A: SONA Posting

Study Name: Effects of Training Environment on prospective memory performance.

Experimenters:

Jinous Mirzaagha, Master's Student, Department of Cognitive Science e-mail: jinousmirzaagha@cmail.carleton.ca

Faculty Supervisor: Kathleen Van Benthem, Senior Research Scientist, Institute of Cognitive Science

Experiment Location: VSIM Building (Building # 38)

In this experiment you will be presented with a VR aircraft environment through a head-mounted display (VR goggles), and with a similar environment, but presented through a 2D visual display on a computer monitor. In neither condition do you actually have to fly the virtual aircraft, however, we will ask you to perform some memory tasks during both flights. The purpose of this study is to learn about how Training Environments help with memory tasks during virtual flight activities. This research was cleared by the Carleton University Research Ethics Board – B (Clearance # 111356) on October, 2019.

Eligibility Requirements: Normal or corrected-to-normal visual acuity, without colour blindness, and normal or corrected-to-normal hearing.

Duration: 1.5 hours

SONA Student Remuneration: SONA participants will receive 1.5% credit for their participation. All participants will be still be compensated (as just described), even if they withdraw from the experiment at any point between signing the informed consent form and the end of the experiment.

Preparation: None

Study Risks: There is a mild and rare risk for simulator sickness with this experiment. Simulator sickness is alleviated quickly by closing your eyes and then having a drink of water, which we will provide if necessary.

Exclusions: Please do not sign up for this study if you have a known history of nausea when playing video games or if you experience motion sickness in real or virtual moving vehicles.

Appendix B: Recruitment poster for non-students

Study Name: Effects of Training Environment on prospective memory performance.

Experimenters:

Jinous Mirzaagha, Master's Student, Department of Cognitive Science e-mail: jinousmirzaagha@cmail.carleton.ca

Faculty Supervisor: Kathleen Van Benthem, Institute of Cognitive Science

Experiment Location: VSIM Building (Building # 38)

Description: In this experiment you will be presented with a VR aircraft environment through a head-mounted display (VR goggles), and with a similar environment, but presented through a 2D visual display on a computer monitor. In neither condition do you actually have to fly the virtual aircraft, however, we will ask you to perform some memory tasks during both flights. The purpose of this study is to learn about how Training Environments help with memory tasks during virtual flight activities. This research was cleared by the Carleton University Research Ethics Board – B (Clearance # 111356) on October, 2019.

Eligibility Requirements: Must be at least 50 years old with normal or corrected-to-normal visual acuity, without colour blindness, and normal or corrected-to-normal hearing.

Duration: 1.5 hours

Student Compensation: Non-student participants will receive complimentary refreshments (value ~$5) for their participation, as well as a parking pass for the duration of the experiment. Participants will still be compensated even if they withdraw from the experiment at any point between signing the informed consent form and the end of the experiment.

Preparation: None

Study Risks: There is a mild and rare risk for simulator sickness with this experiment. Simulator sickness is alleviated quickly by closing your eyes and then having a drink of water, which we will provide if necessary.

Exclusions: Please do not sign up for this study if you have a known history of nausea when playing video games or if you experience motion sickness in real or virtual moving vehicles.

Appendix C: Informed Consent

Project Title: Perception in VR flight simulation

Faculty Sponsor: Dr. Kathleen Van Benthem, Department of Psychology, Carleton University, tel. 520- 2600 x. 8122

The purpose of this informed consent form is to ensure that you understand both the purpose of the study and the nature of your participation. The informed consent must provide you with enough information so that you have the opportunity to determine whether you wish to participate in the study. This research was approved by the Carleton University Research Ethics Board – B (Clearance # 111356) on October, 2019. Please ask the researcher to clarify any concerns that you may have after reading this form.

Research Personnel: In addition to the Faculty Sponsor named above, the following people are involved in this research and may be contacted at any time should you require further information about this study:

Name	**Title**	**Department**	**Email**
Jinous Mirzaagha	Master's Student	Department of Cognitive Science	jinousmirzaagha@cmail.carleton.ca
Kathleen Van Benthem	Postdoctoral Fellow	Institute of Cognitive Science	kathy.vanbenthem@carleton.ca
Sarah Arseneau	Research Assistant	Institute of Cognitive Science	sarah.arseneau3@carleton.ca

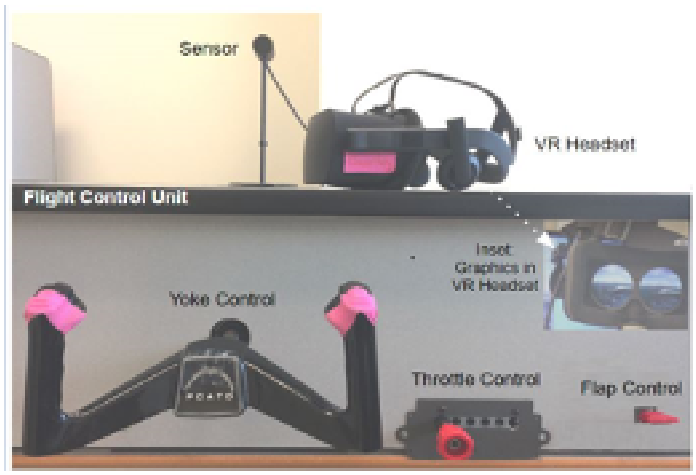
Other Contacts: Should you have any ethical concerns with the study, please contact the REB Chair, Carleton University Research Ethics Board-B (by phone: 613-520-2600 ext. 4085 or by email: ethics@carleton.ca). For all other questions about the study, please contact the researcher.

Purpose: We are interested in the effects mode of training on prospective memory performance in the flight environment.

Task Description: In this experiment you will be presented with approximately 15 minutes of flight training in either a VR (VR goggles, or 2D (computer monitor graphics) environment. The simulated environment will look like a flight route above the City of Las Vegas. Participants will be asked complete some prospective memory tasks when you pass by certain landmarks and/or objects in the environment. Training will include learning about the prospective memory tasks, the cues that will be used that indicate when a task should be completed, and learning how to use the VR and standard response interface (VR response pad) The Test Graphics phase will consist of two conditions that are approximately 15 minutes each, and you will be tested in both VR and 2D environments.

We will follow-up with a questionnaire to ask questions about your experience in flying, and any discomfort you might have experienced during flight.

HTC Vive Headset for VR Virtual keypad to enter information



Flight control unit (the flight path is automated and you do not have to fly the aircraft).

Duration, Locale & Compensation: Testing will take place in the Visualization and Simulation Building on the Carleton Campus and will take approximately 1.5 hours. SONA students will receive 1.5% course credit for your participation and other participants will

receive complimentary refreshments valued at ~ $5.00. You will receive this compensation, even if you choose to withdraw at any point after signing the consent form.

Potential Risks/Discomfort: Some participants may feel a bit nauseous or dizzy, this is a rare and mild risk for this experiment. In our experience this is very rare and if it occurs it subsides quickly by closing your eyes and drinking some water, which we will provide. Please alert the experimenter if you need a break or wish to stop. If participants become uncomfortable at any point please inform the experimenter. If you wish to withdraw you may do so without penalty.

Anonymity/Confidentiality: All data collected in this experiment will be kept strictly confidential and will be anonymized through the assignment of a coded number and securely stored on a local server for a maximum of ten years. This Informed Consent form will be kept for a maximum of ten years before being destroyed. The information provided will be used for research purposes only. You will not be identified by name in any reports produced from this study. Further, the information is made available only to the researchers associated with this experiment. The behavioural data will be archived and kept on a password protected server in the ACE lab for a maximum of 10 years for potential future use in similar projects. The data will only be shared in publications and presentations as averages, and individuals will not be identifiable. The anonymous online study data will be stored and protected by Google Forms TM on their secure servers but may be disclosed via a court order or data breach.

Right to Withdraw/Omit: You have the right to withdraw from this experiment at any time without academic penalty. However, you cannot withdraw after leaving the experimenter session. Your participation in this experiment is completely voluntary, please alert the experimenter if you wish to withdraw.

I have read the above description of the study effects of the presence of virtual hands in the flight environment on flying performance and mental workload. By signing below, this indicates that I agree to participate in the study, and this in no way constitutes a waiver of my rights.

Name: ______________________Signature: _____________________

Date: ______________________ Witness: ______________________

Appendix D: Debriefing form

Study Title: Effects of Training Environment on prospective memory performance.

Research Personnel: In addition to the Faculty Sponsor named above, the following people are involved in this research and may be contacted at any time should you require further information about this study:

Name	Title	Department	Email
Jinous Mirzaagha	Master's Student	Department of Cognitive Science	jinousmirzaagha@cmail.carleton.ca
Kathleen Van Benthem	Postdoctoral Fellow	Institute of Cognitive Science	Kathy.vanbenthem@carleton.ca

Other Contacts: Should you have any ethical concerns with the study, please contact the REB Chair, Carleton University Research Ethics Board-B (by phone: 613-520-2600 ext. 4085 or by email: ethics@carleton.ca). For all other questions about the study, please contact the researcher.

Thank you for participating in this experiment. The results of this study may provide empirical support for how cognitive mechanisms work in VR. The results will benefit future development of VR systems for training and cognitive health screening for general aviation, particularly for individuals with cognitive impairments or older pilots. This study aims to learn more about the requirements in flight training and screening tools, and if there is a relationship between Training Environment and testing environment when asked to perform a challenging ongoing memory task and detect environmental cues that will support performing prospective memory tasks. Research has shown that there are a number of factors that can influence prospective memory performance, such as individual differences (e.g. old vs. Young, experience), or encoding specificity where cues require an adequate trigger formed during the association process that can cue prospective memory if present. For this reason, we are also hoping to learn about how older age might moderate any effects of the features of the task Training Environment on prospective memory. This will help us as we develop a VR assessment and training tool for pilots. For information on prior research in the area of prospective memory please see:

Van Benthem, K., Herdman, C.M., Tolton, R.G., & LeFevre, J. (2015). Prospective memory failures: effects of cue salience, workload, and pilot individual differences. Aerospace Medicine and Human Performance, 86, 366-373.

This research was approved by the Carleton University Research Ethics Board – B (Clearance # 111356) on October, 2019. Should you have any ethical concerns with the study, please contact the REB Chair, Carleton University Research Ethics Board-B (by phone: 613-520-2600 ext. 4085 or by email: ethics@carleton.ca). For all other questions about the study, please contact the researcher. Should you have any other concerns about this study then please contact any of the following individuals:

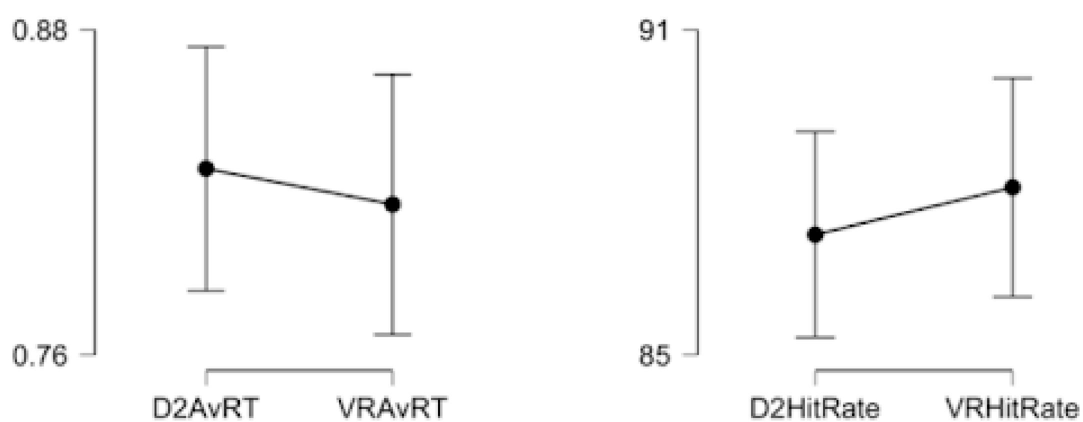
Name	**Title**	**Department**	**Study Role**	**Email**
Jinous Mirzaagha	Master's Student	Cognitive Science	Lead Researcher	jinousmirzaagha@cmail.carleton.ca

Kathy Van Benthem	Postdoctoral Fellow	Institute of Cognitive Science	Academic Supervisor	Kathy.vanbenthem@carleton.ca
Sarah Arseneau	Undergraduate	Institute of Cognitive Science	Research Assistant	sarah.arseneau3@carleton.ca

Appendix E: Auditory Peripheral Detection Task

An auditory peripheral detection task was used as an index of mental workload to confirm that the two Test Graphics conditions (2D versus VR) did not differ significantly in workload. A paired-samples *t-test*, conducted with a sub-sample of participants (n=13), found that average response times and hit rate for the PDT were similar across the two Test Graphics conditions. The PDT analysis indicated that should differences in prospective memory be found across Test Graphics conditions, that this difference was not a result of one condition being associated with higher levels of mental workload i.e., each condition was equally challenging.

As shown in the figures below, there was no significant difference in response times for the 2D (M = 0.82, SD = 0.12) and or VR (M = 0.816, SD = 0.162) conditions; $t(15)$ = 0.204, p = 0.841. Similarly, there was no significant difference in hit rate for the 2D (M = 87, SD = 0.14) and or VR (M = 88, SD = 13.4) conditions; $t(15)$ = -1.23, p = 0.233.



Appendix F: SPSS Syntax for Post-hoc Analysis for Graphics Type and Cue Type Interaction

```
GLM D2_FlapsDownPercent D2_AirspeedCalls_Percent VR_FlapsDownPercent
VR_AirspeedCalls_Percent
  /WSFACTOR=DisplayType 2 Polynomial Cue_Type 2 Polynomial
  /METHOD=SSTYPE(3)
  /PLOT=PROFILE(DisplayType*Cue_Type) TYPE=LINE ERRORBAR=CI
MEANREFERENCE=NO YAXIS=AUTO
  /EMMEANS=TABLES(OVERALL)
  /EMMEANS=TABLES(DisplayType) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(Cue_Type) COMPARE ADJ(BONFERRONI)
  /EMMEANS=TABLES(DisplayType*Cue_Type)
  /PRINT=DESCRIPTIVE ETASQ OPOWER
  /CRITERIA=ALPHA(.05)
  /EMMEANS= TABLES(DisplayType*Cue_Type) COMPARE(Cue_Type)
ADJ(BONFERRONI)  ** Special commands for post-hoc tests
  /EMMEANS= TABLES(DisplayType*Cue_Type) COMPARE(DisplayType)
ADJ(BONFERRONI)   ** Special commands for post-hoc tests
  /WSDESIGN=DisplayType Cue_Type DisplayType*Cue_Type.
```