Examining the Effects of Perceived Telepresence, Interactivity, and Immersion on Pilot Situation Awareness During a Virtual Reality Flight Exercise

by

Cassandra Ommerli

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

Master of Arts

in

Human-Computer Interaction

Carleton University
Ottawa, Ontario

Abstract

Flight simulators that use virtual reality (VR) displays are becoming more prevalent in the aviation domain, enabling situation awareness (SA) training and assessment paradigms that integrate a broad range of aircraft types and flight environments. Research has identified three key components of user psychological experiences during VR exposure that may affect cognitive performance: immersion, telepresence, and interactivity. The primary objective of this study was to validate and quantify the effects of these VR experience constructs on pilot SA at the levels of information processing and comprehension. Effects of age and experience were also explored, as these factors are known to influence achievement of SA. Findings from this research will provide insight into the ways in which pilots' psychological experiences of VR flight simulators affect their cognitive processing. Moreover, the results of this work will inform the design of improved VR systems for the training and assessment of pilot SA.

Acknowledgements

I express my sincere gratitude to Dr. Chris Herdman for the opportunity to complete my graduate studies under his supervision in the Advanced Cognitive Engineering (ACE) Laboratory. I am grateful for not only his ongoing academic guidance and compassionate demeanour towards students, but for his encouragement of my professional goals.

Next, I would like to recognize the tireless contributions of Dr. Kathleen Van Benthem in operation of ACE Lab studies and support of student work. Kathy's mentorship has provided me with not only indispensable knowledge in research design, data collection, and statistical analysis, but has inspired a greater confidence in me as a researcher. Thanks go out also to former postdoctoral fellow Dr. Chunyun Ma, as well as numerous ACE lab students and engineering staff, who have contributed their time to the larger study on pilot cognitive evaluation that serves as the foundation for my own branch of work.

I thank Dr. Audrey Giourard and other Carleton faculty for facilitating critical work opportunities for graduate students through the Collaborative Learning of Usability Experiences (CLUE) program, which has enabled a new career path for me as a user experience (UX) researcher. Thanks go out to my CLUE mentor, Dominira Saul, for setting a precedent in leadership and inspiring a new generation of UX professionals.

Thank you also to Dr. Aminah Robinson Fayek at the University of Alberta for providing me with an exciting academic and professional learning opportunity in working as a technical writer for her engineering R&D program, and for supporting me extensively in pursuit of my own post-graduate education.

Finally, I thank my partner, Andrew, and parents, brother, and friends for encouraging me to pursue greater academic and professional development. Thank you for continuing to push me forward through the more challenging moments!

Table of Contents

List of Tables

List of Figures

List of Appendices

1. Introduction

Flight simulators that use immersive virtual reality (VR) displays are becoming more prevalent in the general aviation training domain. VR systems are useful across all levels of flight operations, as they enable situation awareness (SA) training and assessment paradigms that integrate a wide range of aircraft types and flight environments. SA is regarded as a critical driver of decision making in all complex and dynamic systems, and it is acknowledged as an integral component of successful performance and safety outcomes in aviation. The present research uses Endsley's (1995a, p. 36) three-level model, which articulates SA as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and projection of their status in the near future." Pilot age and experience are also critical to the present discussion, as the cognitive structures that facilitate SA acquisition are thought to be sensitive to age-related changes, such as decreased information processing speed and resource availability (Salthouse, 1991; Bolstad & Hess, 2000). Moreover, the average age of general aviation pilots is increasing, with many pilots remaining active in flight well beyond age 60. In contrast, age effects may be mitigated in situations where an individual is highly experienced, due potentially to the activation of schemata in long-term memory (Hess, 1990). Although researchers in the VR domain have long been concerned with establishing criteria for evaluating user experiences in virtual environments, there is a lack of research addressing the impact of VR attributes on SA during flight. There is also limited discussion on the effects of age on individual's experiences during VR exposure.

Research in VR technology has identified three key components of user psychological experiences in simulated environments that may affect cognitive performance: immersion, telepresence, and interactivity (Mütterlein, 2018). Telepresence refers to the subjective experience of being physically located within a simulated environment (Minsky, 1980). Achievement of telepresence experience is often treated as an indicator of the ecological validity of VR devices, and as a necessary condition for potential transfer of skills and knowledge acquired in virtual training conditions to real-world contexts. Interactivity refers to a sense of agency, or the perceived capacity to influence the form of content in a simulated environment (Steur, 1992). Finally, immersion is characterized as the subjective experience of feeling involved or absorbed in activities conducted in a virtual environment. The present study uses a flow-based definition of immersion. Flow may be described as a psychological state of optimal experience, in which an individual is entirely absorbed or immersed in an activity (Csikszentmihalyi, 1990).

The primary objective of the present research was to validate and quantify the effects of telepresence, interactivity, and immersion on pilot SA at the first and second level (i.e. information processing and comprehension, respectively). It was hypothesized that strong experiences of telepresence, immersion, and interactivity would help to streamline pilot attention to relevant stimuli, while blocking out irrelevant stimuli from the external environment. As a secondary line of inquiry, the present research will evaluate correlations between age and an individuals' perceived telepresence, interactivity, and immersion.

This thesis is structured as follows. Section I introduces the present study. Section II provides an overview of literature, which informed this work. Section III summarizes the hypotheses assessed by this study, and the steps undertaken to validate a model of relations among the aforementioned VR and SA factors. Section IV gives an overview of the study methodology, including participant characteristics, materials, and experimental procedure. Section V covers results of a quantitative analysis performed on the data collected during the experimental exercise. Section VI presents a deeper analysis of results observed during the present study, in addition to identifying additional opportunities for research. Finally, Section VII concludes the paper, addressing limitations related to the current study, as well as the main findings and contributions.

Findings from this research will help to advance the discussion concerning the ways in which pilots' psychological experiences of VR flight simulators are related to their cognitive processing. Moreover, the results of this work will inform improvements in the design of immersive VR systems for the purposes of the training and assessment of pilot SA.

2. Literature Review

2.1. Virtual Reality Flight Simulators

2.1.1. Practical Considerations

In the aviation domain, VR flight simulators are becoming more sophisticated, enabling representation of a wider range of aircraft types, training paradigms, and environmental conditions. Moreover, VR technology is increasingly being used for training and evaluation in contexts such as operation of military vehicles, maintenance of mechanical systems, nuclear power plant operation, and surgical procedures (Gorman, Meier, & Krummel, 1999). All discussion of VR technology throughout this thesis refers to immersive VR displays, such as the Oculus Rift headset. Immersive VR technology is characterized by the manipulation of sensory input to create conditions in which a user feels present within a simulation and able to physically interact with attributes of the environment (Psotka, 1995). At the most practical level, VR provides distinct advantages in portability and cost over conventional high-fidelity pilot training simulators. Traditional flight simulators are generally constructed using authentic physical instrumentation and system layouts, which are expensive to purchase and maintain, especially considering that each aircraft requires specific components (Dörr, Schiefele, & Kubatt, 2000). In addition, simulators are often large in size and can be difficult to dismantle and reconfigure, all of which places a large demand on operators in ensuring the necessary facilities and support personnel. Despite the positive outlook around VR technology in aviation, little is known about how attributes of these systems impact pilot cognitive processes, such as situation awareness.

2.2. Situation Awareness and Pilot Performance in General Aviation

General aviation (GA) pilots experience substantially more accidents, including crashes, injuries, and fatalities, than do those in the commercial or military sectors. The Transportation Safety Board of Canada (2018) reported that of 240 aviation accidents recorded in 2017, 145 (60%) came from privately operated or “other type” aircraft, the latter of which include organizations such as flight clubs and schools. A 2012 study conducted by the United States Government Accountability Office (GAO), which reviewed patterns in both fatal and non-fatal general aviation accidents occurring between 1999 and 2011, found that single-engine piston airplanes flying personal operations were involved in the majority of incidents. In addition, the study identified “human error” and “loss of control” as common components in most accidents (GAO, 2012). While human error in GA flight is often characterized as a failure in decision making, Endsley (1995b) suggests that a large proportion of these issues are actually errors in situation awareness (SA), whereby the problems lie in a pilot’s perception and/or comprehension of the situation, not in his or her response to it. Such errors are different in nature and require distinctive remedial action.

2.2.1. Endsley's Model of Situation Awareness

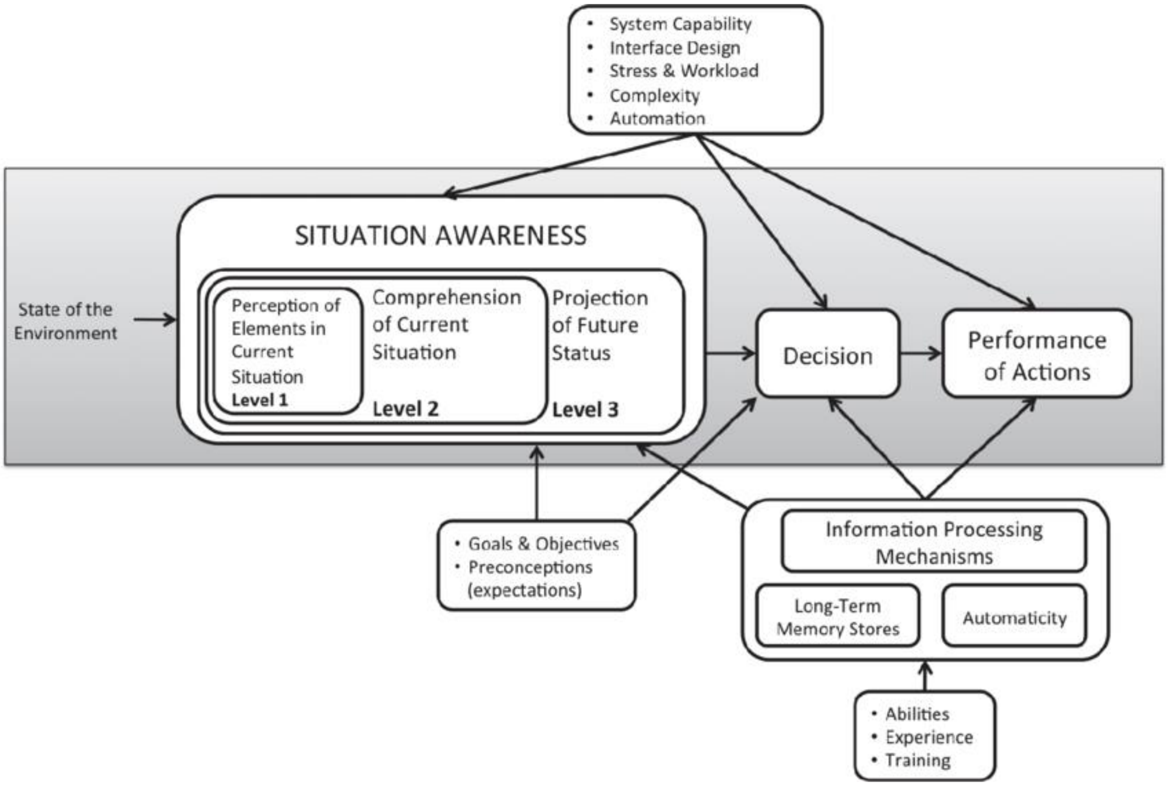


Figure 1. Three-level model of SA from Endsley (1995a), including relationships to underlying cognitive processes, goals and expectations, and system attributes.

SA is regarded as a critical driver of decision making in all complex and dynamic systems involving human operators. SA is a well-studied aspect of pilot cognition and is acknowledged as an integral component of successful performance and safety outcomes in aviation. One of the most widely-cited definitions comes from Endsley (1995a, pp.36), who defines SA as a dynamic state involving "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and projection of their status in the near future." Furthermore, the present research defines SA using the three-level mechanistic model developed by Endsley (1995a), the components of which are depicted above in Figure 1 and discussed in detail below.

2.2.1.1. Level I Situation Awareness

The first level of SA involves perception of the “status, attributes, and dynamics of relevant elements” in one’s environment (Endsley & Jones, 2012). The relevant elements and the type of psychological information conveyed to the agent will vary depending on the environmental context. In GA, a pilot is trained to interpret a range of system, flight, and navigational data being communicated to them via cockpit instruments. However, an experienced pilot will also be attuned to subtle sensory cues, such as changes in the pitch of their engine. In any complex system, pilots must continuously filter large quantities of competing signals, making it difficult to identify all relevant SA Level I data. Endsley and Jones (2012) suggest that approximately 75% of pilot SA errors can be attributed to pilots not perceiving essential information, with nearly 40% of those errors coming from instances where system limitations impeded adequate communication of information.

2.2.1.2. Level II Situation Awareness

SA at the second level involves distillation of large quantities of psychological data into bits of relevant of information, which an individual then interprets for significance with regard to their current goals and environmental conditions. Endsley and Jones (2012) note that approximately 19% of SA errors in aviation involve Level II SA. In such instances, individuals correctly perceive sensory data, but fail to comprehend the significance of it. Given the cognitive demand posed by data integration, pilots with more advanced mental models corresponding to a situation will possess an advantage in their ability to develop Level II SA than do those lacking a comparable knowledge base.

2.2.1.3. Level III Situation Awareness

Finally, Level III SA involves the ability of a pilot to predict future conditions on the basis of one's understanding of current conditions. Level III SA is thus dependent on one's ability to demonstrate adequate comprehension of a situation (i.e., Level II SA). Moreover, the pilot must already possess a sufficiently developed mental model of flying in order to accurately appraise the progression of events. Barring errors involving shortcomings in Level I or II SA, Level III SA errors may be attributed to information processing overload, lack of domain knowledge, or an over-projection of trends (i.e., making a prediction on the basis of an assumption that certain conditions will hold stable) (Endsley & Jones, 2012). Endsley and Jones note that only approximately 6% of errors in SA can be categorized under Level III SA, further suggesting that this low number might be attributed to the high cognitive demand already placed on an individual to complete the prerequisite information processing at levels I and II. Bolstad et al. (2010) maintain that good SA is commonly associated with high levels of flight experience, citing Prince and Salas (1997), who compared awareness of SA cues across GA pilots (possessed an average of ~720 hours of flight experience), commercial pilots (~6036 hours) and check pilots (i.e., individuals who evaluate the safety and proficiency of commercial pilots) (~12,370 hours). Findings from Prince and Salas showed that the GA pilots were mostly only aware of direct temporal and spatial Level I SA cues. In contrast, airline pilots fared better in also actively identifying and attending to Level II SA cues, while only check pilots were widely able to develop an advanced projection future conditions and manage complex events accordingly (Level III SA cues).

2.2.2. Temporal Demands

Given that a key characteristic of dynamic systems is the continuous temporal evolution of the system state, time is another major component in Endsley's model of SA. Time is especially important in the context of comprehension and projection, as successful SA at these levels depends on the capacity of an individual to understand the rate at which conditions are changing. For example, in order for a pilot to avoid an accident situation after recognition of the issue, they must also have some level of awareness regarding the immediacy with which corrective action is required. Since the state of a system is continuously changing, SA should not be thought of as a static state, but instead as involving stages of information processing. An up-to-date projection of the situation requires pilots to continuously take in and parse relevant data to update their SA.

2.2.3. Relationship to Other Cognitive Processes

SA occurs as the product of a number of underlying cognitive processes. For example, limitations in attention and working memory are thought to be primary bottlenecks for SA, in that they dictate the amount of unrelated data that can be identified and processed by an individual at any given moment. Working memory is a limited-capacity cognitive storage system that also enables the processing of information. New information is combined with existing chunks in working memory to generate an updated picture of one's environment, while information not being engaged rapidly decays. One of the most established models of working memory comes from Baddeley and Hitch (1974) and Baddeley (2000), which describes a three-component structure. The primary component is the central executive, which is responsible for directing attention to relevant information, while suppressing out irrelevant information. The central executive also coordinates the activities of three “slave systems”: the phonological loop, which maintains verbal information, the visuospatial sketchpad, which maintains visual and spatial information, and the episodic buffer, which integrates information from the aforementioned systems.

In contrast, limitations in SA can be overcome by strengthening other cognitive mechanisms. Long-term memory structures, such as mental models (complex structures that individuals use to model the behaviour of systems) and schema (states of the mental model that organize categories of information and facilitate efficiencies in information processing), support SA by helping an individual focus on relevant information and form expectations for the task environment. Endlsey and Jones (2012) underscore mental models as “key enablers” of complex SA processes (i.e., comprehension and projection).

2.2.4. Age and Experience

The average age of general aviation pilots is increasing, and it is not uncommon for individuals to remain active in flight well beyond age 60. A 2018 study by the Federal Aviation Administration found that within a population of 633,316 US citizens with active pilot certificates, 147,984 (23%) were 60 years of age and older. Age-related cognitive decline is frequently associated with reduced information processing resources (Salthouse, 1991). The cognitive structures that facilitate SA acquisition are thought to be sensitive to such changes. However, the discussion is complicated by considerable variability in the incidence and onset of decline among individuals (Bolstad & Hess, 2000). Bolstad and Hess (2000) break down potential effects of cognitive decline across each of the three levels of SA. At the first level, age-related decline in attentional and processing resources may limit the amount of information that enters working memory, the efficiency of an individual's retrieval processes, and the complexity of information that can be handled at any given moment. In turn, these changes may reduce a pilot's ability to generate an accurate mental model of their environment, the latter of which enables second- and third-level SA (i.e. comprehension and projection of future states, respectively). Moreover, older adults may demonstrate weaker source memory than younger adults, which entails difficulty in encoding contextual information. Parsing informational context is essential for retention of information and is foundational for achieving higher levels of SA.

In contrast, experience may modify the extent to which age effects are observed in different contexts (Bolstad & Hess, 2000). There is evidence to suggest that the impact of age-related cognitive decline is reduced in situations in which an individual is highly experienced, due potentially to the activation of schemata in long-term memory (Hess, 1990). Research from Hess (1990) suggests that schema use and activation appear to be minimally affected by age, and thus serve as critical processing resources for aged individuals.

2.3. Users' Psychological Experiences in Virtual Reality

Researchers of VR technology have identified three key attributes of users' psychological experiences during VR exposure: immersion, presence/telepresence, and interactivity. However, there remains contention in the definition of these experiences within the literature, and limited research has been dedicated to understanding how such experiences relate to and influence one another (Mütterlein, 2018). Therefore, the present research is intended to address gaps in the literature by providing validation for the aforementioned three constructs of users' psychological experiences of VR environments, as well as assessing relationships among these factors.

2.3.1. Presence/Telepresence

A critical point of discussion in the VR literature is the user's experience of presence or telepresence. Telepresence is a term occasionally used in VR research to distinguish states of presence that are achieved through interactions with an environment presented via a technological medium (Steur, 1992). Telepresence is used throughout the rest of this document to denote presence experiences facilitated through the use of VR technology. Telepresence is often treated as an indicator of the ecological validity of VR devices, and it is regarded as a necessary condition for the transfer of skills and knowledge to real-world contexts (Regenbrecht, Schubert, & Friedman, 1998; Mestre, 2006). In general, telepresence is defined as the perception or subjective experience of physically being there within a simulated environment (Minsky, 1980). Moreover, the International Society for Presence Research (2001) suggests that telepresence occurs when an individual's perception fails to acknowledge the interaction with technology, which enables the sensation of feeling that one is in a location different from their actual location in the physical world. Similarly, Heeter (1992) describes telepresence as an ongoing experience of discerning and validating one's existence as a component of, but separate entity within the virtual world. Witmer and Singer (1998) suggest that the level of telepresence an individual experiences is dependent on their ability to shift attentional resources from their external environment to the virtual one. The authors refer to work by McGreevy (1992), who discusses the roles of continuity, coherence, and connectedness in stimulus flow for facilitating telepresence.

Telepresence is a meaningful psychological phenomenon (Schubert, Friedman, & Regenbrecht, 2001); it observable in individuals' physical responses to simulated stimuli. Regenbrecht, Schubert, and Friedman (1998) evaluated participants' emotional arousal when placed before a virtual cliff and concluded that perceived level of presence had a significant positive impact on reported fear. Sanchez-Vives and Slater (2005) argue that telepresence might be seen as consciousness in a virtual environment, where external sensory data is successfully substituted by simulated sensory data in a way that allows an individual to respond via normal cognitive mechanisms. Sanchez-Vives and Slater (2005) refer to the work of Damasio (1998), who characterizes core consciousness as a transient and ongoing process, whereby a sense of self and awareness are automatically generated via short-term memory. What is crucial to this core consciousness is both the automatic sense that an individual is involved in the perception of a stimulus, as well as the sense that they possess the agency to act (or not act) upon that stimulus. Sanchez-Vives and Slater (2005) suggest that telepresence occurs when these qualities are realized in a virtual environment.

Experiences of presence can arise during interactions with content in many different media formats, such as radio, television, and books. For example, an individual reading a book may imagine themselves to be present in the world depicted by the text and create a mental model of their bodies and the actions they carry out (Biocca, 1997; Wirth et al. 2003). While engagement of an individual's sensory channels through visual, auditory, haptic, or proprioceptive feedback may enhance experiences of telepresence, they do not appear to be necessary conditions for their occurrence (Wirth et al., 2003). Moreover, Schubert et al. (2001) suggest that it would be misleading to assume a one-to-one relationship between the stimuli produced by the “immersive” attributes of a system and an individual’s experience of telepresence. The authors maintain that the relationship between stimuli perception and presence is mediated by cognitive processes, which include the suppression of conflicting stimuli from the external environment and allocation of attention to stimuli from the media environment.

2.3.2. Interactivity

The next VR psychological experience is interactivity, which refers to a sense of agency, or perceived capacity to influence the form of content in a simulated environment (Steur, 1992). As compared to presence and immersion, interactivity is less established as a psychological experience construct in the VR literature. Kiousis (2002) examines the varied nature of the way the interactivity of a technological medium is discussed and evaluated in academic research, noting that accounts tend to fall into two major camps depending on whether they focus on the capacity of the technology to facilitate interpersonal communications or not. Within these camps, definitions of interactivity focus on one of three areas: the technological attributes of a system that facilitate interactivity, the context or environment of interactivity, or the perceptions of the user. The present research will focus on user perceptions. For example, McMillan and Hwang (2002) investigated measures of perceived interactivity. The authors identified direction of communication (direct two-way communication vs. indirect exchange), tools and features, and system responsiveness as integral to users’ perception of control over the form of content.

Perceived interactivity seems to be closely related to experiences of telepresence. Though interactivity is not treated as a separate construct in the following example, Wirth et al. (2003) maintain that an individual's perception of the possibility to act within a mediated space is a secondary dimension of telepresence, occurring alongside the perception of being physically situated in an environment conveyed by some medium. A sense of agency and capacity for action appears to be crucial in enabling an individual to experience oneself as existing within a virtual space. Schubert, Friedmann, and Regenbrecht (1998) and Schubert et al. (2001) evaluate telepresence through the lens of embodied cognition, whereby telepresence is seen as emerging through the ongoing construction of a "spatial-functional" mental model of a simulated environment. Schubert et al. (2001) suggest that two crucial cognitive processes facilitate this model, namely the suppression of conflicting stimuli (discussed above), and the continuous representation of bodily action in the virtual world. The authors suggest that our cognitive representations of an environment comprise "patterns of possible action", which are meshed with patterns of action based on memory. Meaning is then created through the interactions between our bodies and attributes of our environment, and the ongoing negotiation of those patterns of action. Following an analysis of telepresence experiences during VR gameplay, Schubert et al. (1998) found that participants' experiences of telepresence were directly impacted by their perceptions regarding the possibility for action in a spatial environment, as well as the occurrence of dramatic events that structured interactions within the game.

2.3.3. Immersion as a Facet of System Capabilities

An enduring debate in VR research concerns whether immersion should be investigated through objective measures as a physiological experience facilitated by the technological capabilities of a system, versus as a subjective psychological state experienced by the user. Slater and Wilbur (1997) define immersion as the extent to which the system is capable of delivering to the user an "inclusive, extensive, surrounding, and vivid illusion of reality". The authors frame their discussion of immersion objectively in terms of a quantifiable set of system capabilities; for example, the inclusiveness aspect refers to the capacity of the system to block out competing sensory input from the user's external environment, while vividness is concerned with the resolution and fidelity of displays. In contrast, Mütterlein (2018) suggests that while immersion certainly appears to be constrained by the technological capabilities of the system, it also appears that users will experience varying levels of "immersiveness" depending on the type of VR content they are exposed to, even if the same technology is employed in both cases.

2.3.4. Immersion as a Flow-based, Psychological Experience

From the opposing view, Witmer and Singer (1998) offer a psychological definition of immersion often cited in the VR literature, which identifies a sense of being "enveloped by" or "included in" the virtual environment. The authors identify a number of factors affecting immersion, such as a sense of isolation from the physical or external environment, the perception that one is included within in the environment, the perception of motion, and a feeling of being able to move and control attributes of the environment in a natural manner. However, the language used in the aforementioned definition appears to overlap with discussions of telepresence and interactivity. Despite the empirical support for telepresence as a distinctive mental state with implications on individuals' performance and learning transfer in VR environments, do we have good reason to treat immersion as an independent psychological factor? Mütterlein and Hess

(2018) and Mütterlein (2018) maintain that immersion and telepresence represent conceptually distinct psychological experiences. The authors suggest value in the use of flow-based measures of psychological immersion to clarify ambiguities between immersion and telepresence definitions. Here, immersion picks out the subjective experience of feeling involved or absorbed in activities conducted in a virtual environment.

Research by Csikszentmihalyi (1990) identifies flow as a psychological state of optimal experience, whereby an individual is entirely absorbed or immersed in an activity. Flow is sometimes colloquially described as the state of “being in the zone”. Activities that induce this state are thought to be autotelic in nature or motivating in and of themselves. Flow is typically measured as a subjective experience, and has been described across nine dimensions: intrinsic motivation to engage in an activity, strong concentration, a sense of control over one’s actions, loss of awareness of the passage of time, limited self-reflection, a clear sense of one’s objectives, balance between an individual’s capabilities and the skill demands of an activity, merging of action and awareness, and processing of immediate and unambiguous feedback (Csikszentmihalyi, 1990; Michailidis, Balaguer-Ballester, & He, 2018). Rheinberg, Engeser, and Vollmeyer (2003) developed the Flow Short Scale to assess flow-based experiences occurring in the course of human-computer interactions. Following a factor analysis of results from the questionnaire, the authors found that flow as a construct could be separated into two key factors: absorption by activity and fluency of performance. Absorption by activity refers to a sense of feeling fully involved in some task, whereby a user’s sense of the passage of time and awareness of their own behaviour falls away. It is suggested that flow is a neurologically meaningful state, where brain structures enabling self-reflection are inhibited during instances of high task demand (Goldberg, Harel, & Malach, 2006; Engeser, 2008). In contrast, fluency of performance entails a strong level of focus and concentration, coupled with clear goals and a marked sense of control over one's actions (Weibel et al., 2008). These flow-based immersion constructs have been employed in the present research. Validation of flow-based measures of immersion is important insofar as researchers are interested in assessing the suitability of VR systems for learning and knowledge transfer. Flow-based experiences, especially pronounced concentration and absorption, are thought to be components of meaningful learning in terms of depth of cognitive processing and performance (Shernoff and Csikszentmihalyi, 2009). In turn, it is crucial to understand how perceived telepresence and interactivity impact an individual's immersion in activities performed in a VR setting.

2.3.5. Relationships Among Virtual Reality Psychological Experience Constructs

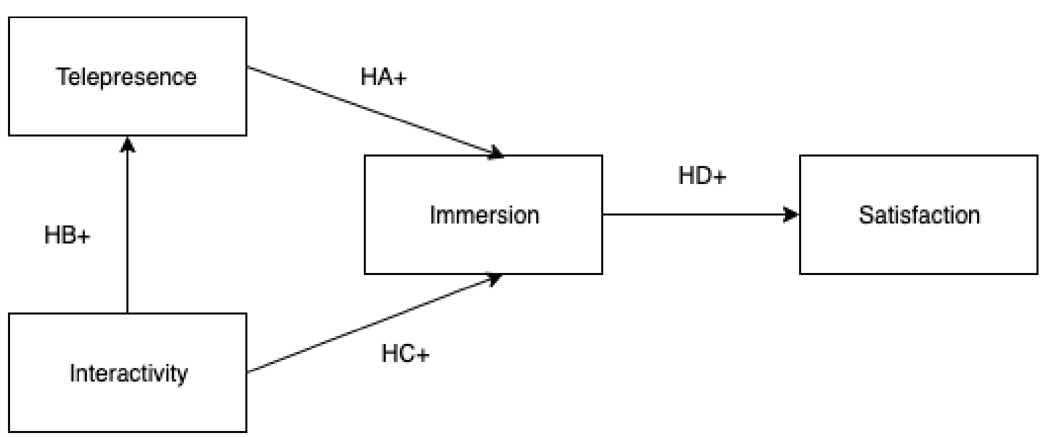


Figure 2. Model of relationships among telepresence, immersion, interactivity, and satisfaction from Mütterlein (2018).

The literature has yet to establish a unified account of the relationships between users' psychological experiences in VR. Emphasizing the interdependent nature of telepresence, immersion, and interactivity, Mütterlein (2018) proposes the model shown above in Figure 2, where HA+, HB+, HC+, and HD+ all represent hypotheses of direct positive relationships between VR experience constructs, which were supported by the author's analysis. In this study, the author administered a 16-item post-test questionnaire regarding subjective experiences of telepresence, interactivity, and immersion to a sample of 294 visitors following a 30-minute exercise conducted using an HTC Vive VR headset. Results from a factor analysis showed all items predictably loaded onto the proposed constructs, including both dimensions of flow (i.e. absorption by activity and fluence of performance) described in the previous section. Further assessment using a partial-least squares approach provided statistical support for the aforementioned relationships.

Telepresence was found to have a direct positive impact on immersion, suggesting that one must first feel as though they are sufficiently present in a virtual environment to experience a sense of immersion in tasks they are performing. In addition, interactivity was also demonstrated to have a direct positive impact on both immersion and telepresence. This finding suggests that the user experiencing feedback from interactions and establishing a sense of control over content, as opposed to passively consuming it, was crucial for achieving a heightened sensation of being present in the virtual world, as well as high levels of absorption and concentration. Though less relevant to the current discussion, participants' level of immersion was shown to have a significant positive impact on their overall satisfaction during use of the system, which represents an important consideration in ensuring that design decisions around VR devices remain empathetic to the experiences of the user. The present research will adopt similar analytical methods in an attempt to validate the proposed relationships between interactivity, immersion, and telepresence, while also investigating potential effects on pilots' levels of SA during a flight exercise in a VR simulator.

2.3.6. Effects of Virtual Reality Psychological Experiences on Situation Awareness

Research on design for immersive VR technology has long been focused on the attributes of a system that facilitate a sense of being physically present in a simulated environment (Riley, Kaber, & Draper, 2004). Telepresence is often treated as an indicator of the "ecological validity" of VR devices, and it is thought to be a necessary condition for potential transfer of skills and knowledge acquired in virtual training conditions to real-world contexts (Mestre, 2006). Buttussi and Chittaro (2018) compared participants' knowledge retention and reported self-efficacy, engagement, and telepresence after completing a serious game (i.e. a game intended for training purposes rather than entertainment) on aviation safety procedures using different VR displays. Three VR displays were evaluated: a desktop VR display, an HMD with a narrow field of view and a tracker with 3 degrees of freedom (DOF), and an HMD with a wide field of view and a tracker with 6 DOF. The authors found that participants reported significantly higher telepresence and engagement when comparing results from the desktop display with the 6-DOF tracker HMD. Though knowledge retention after two weeks and reported self-efficacy did not differ significantly across display type, the authors suggest that the increased engagement associated with the higher-fidelity display may make individuals more likely to repeat the exercise. Johnson-Glenberg (2018) suggests that immersive VR environments offer two profound affordances that have serious implications for user learning and skill transfer. An affordance may be characterized as a design attribute, which signals a possible action that is readily perceived by an agent (Norman, 1988). Johnson-Glenberg identifies the first affordance as the feeling of telepresence, and the second as the sense of agency stemming from the manipulation of simulated content (i.e., interactivity). It is thought that experiences of telepresence contribute to a heightened limbic state, which positively affects an individual's level of attention and engagement (i.e. immersion). In contrast, highly "agentic" learning may positively impact an individual's ability to recall elements of declarative knowledge, such as attributes of spatial layout. For these reasons, VR technology seems well-poised to offer a robust flight training environment for pilots that is conducive for building long-term memory structures (i.e. mental models and schema), which support improved SA outcomes in real-world flight.

Moreover, it has been suggested that if a user experiences a sensation that they are directly integrated within the virtual environment, as opposed to interacting with it through some medium, they may achieve enhanced task performance (Bystrom, Barfield, & Hendrix, 1999). Though performance is generally thought to improve as SA improves, strong or weak SA does not guarantee corresponding performance outcomes (Endsley, 1990; Jung, Jo, & Myung, 2008). In an evaluation of mission effectiveness among teams of pilots, Endsley (1990) found that pilots with strong SA were only able to achieve good performance in certain conditions, namely when the rules of engagement permitted certain courses of action, when pilots had the right tactical capabilities to take advantage of their knowledge, and when the right decision was made for the situation (i.e. on the basis of SA). Moreover, pilots with low SA were able improve their performance through behavioural modification. For example, individuals may perform better in contexts where their knowledge is inadequate by choosing more conservative courses of action.

A number of studies have reported significant positive correlations between SA and subjective experiences of telepresence. Prothero et al. (1995) suggest that telepresence and SA are closely related concepts, in that both require an individual to have awareness of “self-orientation and self-location” with regard to their environment. He et al. (2018) found that enhancing the vividness of a VR simulation, as well as its level of interactivity, helped to increase participants’ SA. The authors suggest that providing rich detail in design facilitates continuity in the believability of the experience, which helps users avoid “breaks in presence” that might hinder their awareness. Perceived telepresence, including effective suppression of external stimuli, may be a necessary precondition for an individual to achieve normal SA in a virtual environment. In addition, Jung et al. (2008) report a strong positive correlation between SA and presence, and between presence and performance on a combat-based flight simulator game. However, it should be noted that the authors performed their evaluation on a desktop VR system as opposed to an immersive VR system, such as a head-mounted display. Lapataned (2006) also reports a positive association between presence and SA, though no association was observed between SA and performance and between presence and performance.

Less is known about the specific effects of perceived interactivity and immersion on SA in VR environments. Interactivity is likely to be impacted by attributes, such as the speed in which user input is represented in the environment, the range of possible actions available to the user, and the predictability of the responses generated by user inputs (Reed & Saleem, 2017). It is possible that users who experience strong levels of perceived interactivity will achieve better SA, as high levels of interactivity suggest the user is highly focused on elements of the virtual environment and effectively filtering out external noise stimuli. Moreover, interactivity appears to play a crucial role in enabling telepresence. Insofar as a positive relationship exists between telepresence and SA, interactivity may also have an indirect positive impact through telepresence experience.

The current study employs a flow-based conceptualization of immersion. Flow is thought to be a significant predictor of educational performance outcomes in the context of attention- and memory-demanding activities, such as exam writing (Engeser et al., 2006; and Schüler, 2007). It is possible that achievement of marked concentration and task absorption during the experimental exercise will result in improved SA outcomes by limiting attention placed on unrelated stimuli. Participants that report strong experiences in all three areas may thus demonstrate stronger SA, so long as the stimuli created by the virtual environment that enable telepresence, immersion, and interactivity do not create competition for an individual's attention and memory resources.

2.3.7. Effects of Age on Virtual Reality Psychological Experiences

The VR literature currently offers a limited body of research directly exploring the effects of age on users' psychological experiences in VR. Research by Bangay and Preston (1998) exploring factors affecting individual's experiences in VR environments identified age as having a negative effect on participants' sense of telepresence. The authors note that participants in their oldest age bracket (35-45 years) provided lower scores for telepresence, but higher scores for sense of control over the environment. In contrast, Corriveau Lecavalier et al. (2018) compared episodic memory performance of participants over and under the age of 57 across VR and real-life test conditions. The authors also looked at participants' reported telepresence, motivation, and cybersickness and found there to be no significant difference in perceived telepresence across age groups. Interestingly, memory performance in the VR test condition was shown to be positively correlated with performance in traditional conditions for both age groups, suggesting the validity of VR devices in cognitive assessment.

Age may or may not have a relation to participants' reported immersion in a simulated environment. Payne et al. (2011) investigated flow among adults between 60 and 94 years old and found that an individual's ability to achieve a flow state is not negatively compromised by age. Instead, more cognitively demanding activities (e.g. working or participating in an educational/classroom activity) were shown to result in higher levels of flow among individuals with greater fluid intelligence, while lower levels of flow were observed among those demonstrating low levels of fluid intelligence. The pattern was reversed in the context of activities presenting a low cognitive load (e.g. socializing or exercising). The present research will address gaps in the literature by evaluating correlations between age and an individuals' perceived telepresence, interactivity, and immersion.

3. Hypotheses and Model Building Process

3.1. Main Hypotheses

The primary purpose of the present research was to model the direct and indirect effects of three factors of pilots' psychological experiences during VR exposure, interactivity, telepresence, and immersion, on their ability to achieve SA at the first and second level. The first level of SA entails perception of relevant information, while the second level concerns comprehension of the significance of that information. Three main hypotheses were investigated.

3.1.1. Telepresence has a direct positive effect on both SA Level I (H1a) and SA Level II (H1b)

A number of studies exploring the impact of VR attributes on participant SA report a strong, positive association between telepresence and SA (Laptaned, 2006; Jung et al., 2008; and He et al., 2018). Perceived telepresence, or the experience of realizing oneself as physically existing within a simulated environment, may be a necessary precondition for an individual to achieve normal SA when using a VR system.

3.1.2. Interactivity has a direct, positive effect on both SA Level I (H2a) and SA Level II (H2b)

Telepresence appears to be closely related to interactivity. The perceived capacity for action may be an integral factor in enabling an individual to experience the sensation of physical presence in VR conditions. It is possible that users who experience strong levels of perceived interactivity will achieve better SA, as high levels of interactivity suggest that the user is focused on elements of the virtual environment and effectively filtering out external noise stimuli.

3.1.3. Immersion has a direct, positive effect on both SA Level I (H3a) and SA Level II (H3b)

The present study uses a flow-based definition of immersion, which refers to a psychological state of optimal experience, where an individual is entirely absorbed in an activity (Csikszentmihalyi, 1990). Flow has been observed to be a significant predictor of performance in the context of activities demanding a high level of attention and memory (Engeser et al., 2006; and Schüler, 2007). It was expected that participants who achieved stronger flow states would be more likely to demonstrate better levels of SA I and II in the experimental exercise.

3.2. Validation of Factors

3.2.1. Validation of Virtual Reality Psychological Experience Factors

The first step in the model-building process involved validation of factors for telepresence, interactivity, and immersion. Moreover, relationships between factors were investigated according to the model proposed by Mütterlein (2018). The author suggests that interactivity has a direct positive effect on both telepresence and immersion, and that telepresence positively affects immersion.

3.2.2 Validation of Experience Factor

Pilot experience is an important factor in models of pilot performance, as higher levels of experience are expected to be associated with better performance. Including experience as a factor will permit ecological validity testing for other model factors, such as SA. Experience was constructed as formative (composite) variable, where multiple variables, such as license level and total hours flown, each contribute unique aspects of expertise to the factor.

3.2.3. Validation of Situation Awareness Factors

The final step in the measurement model validation phase involved constructing the factors for SA Levels I and II. The present research uses Endsley's (1995a) model of SA and the Situation Awareness Global Assessment Technique (1995c) (see section 4.3.4 for more information), which are well-validated and widely cited in the SA literature. Given the reliability of this model and technique, a direct positive relationship between SA Level I and SA Level II was expected. Additionally, the cognitive processes that facilitate SA acquisition are thought to be sensitive to age-related decline. As a result, the present work expected age to be negatively correlated with both SA Level I and SA Level II. However, there is evidence to suggest that the impact of age-related cognitive changes is reduced in situations in which an individual is highly experienced (Bolstad & Hess, 2000). Experience was expected to be positively correlated SA Level I and SA Level II, thus moderating the relationships between age and SA at both levels.

3.3. Supplementary Hypotheses

As a secondary focus, the current study examined effects of age on the proposed VR psychological experience factors. Investigating these relationships also helped to identify any peripheral moderating influences on the relationships between the VR psychological experience factors and SA. Three supplementary hypotheses were examined.

3.3.1. Age has a direct, positive effect on interactivity (H4); age has a direct, negative effect on telepresence (H5); and age does not have a direct effect on immersion (H6)

A 1998 study by Bangay and Preston exploring factors affecting individual's experiences in VR environments identified age as having a negative effect on participants' sense of telepresence. In contrast, participants in their oldest age bracket (35-45 years) reported higher scores for sense of control over their environment (i.e. interactivity). Unlike telepresence and interactivity, an individual's ability to achieve a flow state is not thought to be impacted by age.

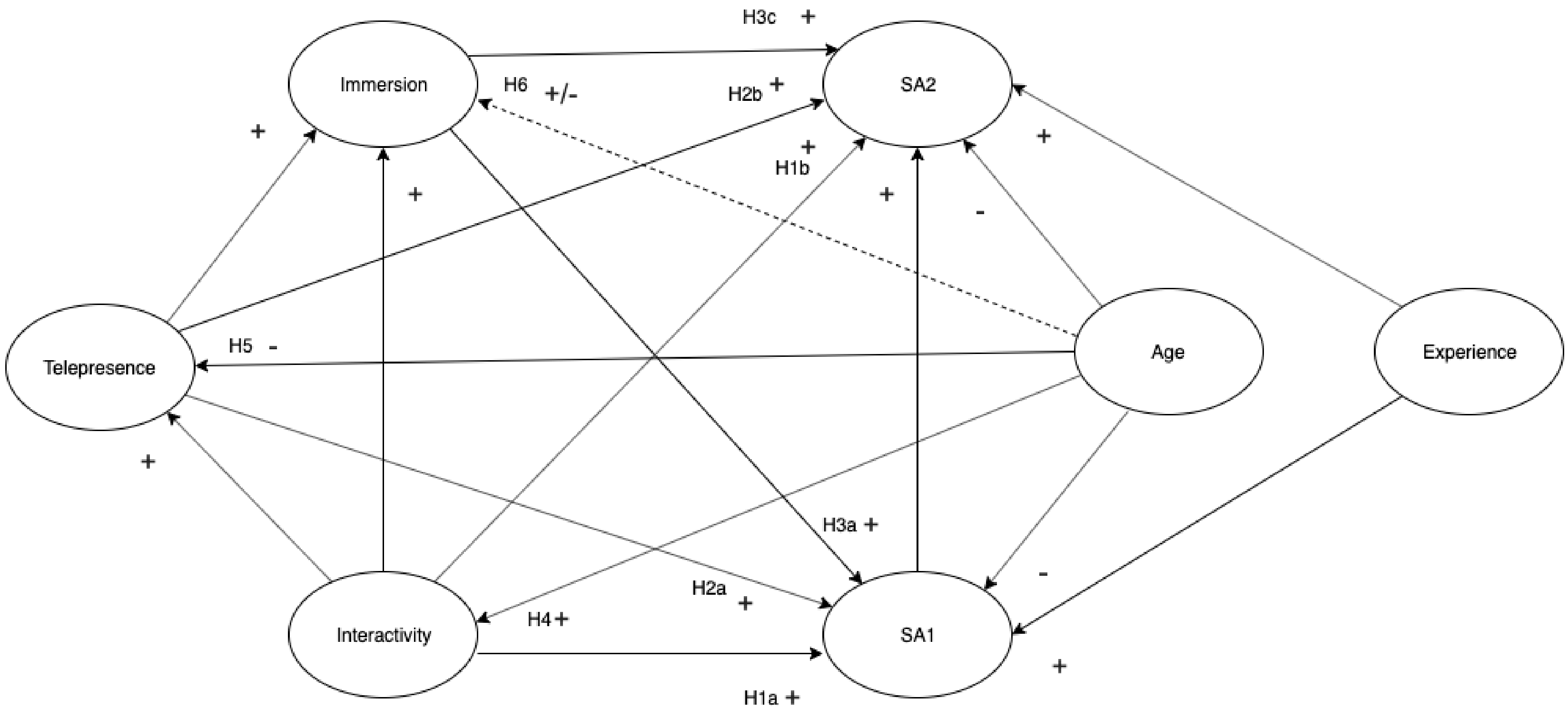
3.4. Model of Psychological and Pilot Factors Influencing Situation Awareness in Virtual Reality

Table 1 provides a summary of the study's main and supplementary hypotheses, as well as the steps of the model-building process. Figure 3 illustrates the combined model of all relationships that were evaluated in the context of the present work.

Table 1.

Summary of All Primary and Supplementary Hypotheses, As Well As Intermediate Validation Steps in Creation of the Combined Model

Hypotheses and Model Building Processes
Relationships between VR psychological experience factors and SA
H1a: Interactivity has a direct, positive effect on SA Level I
H1b: Interactivity has a direct, positive effect on SA Level II
H2a: Telepresence has a direct, positive effect on SA Level I
H2b: Telepresence has a direct, positive effect on SA Level II
H3a: Immersion has a direct, positive effect on SA Level I
H3b: Immersion has a direct, positive effect on SA Level II
Steps:
(one) Validation of VR psychological experience factors
Interactivity has a direct, positive effect on telepresence
Interactivity has a direct, positive effect on immersion
Telepresence has a direct, positive effect on immersion
(two) Validation of SA factors
SA Level I has a direct, positive effect on SA Level II
Age has a direct, negative effect on SA Level I
Age has a direct, negative effect on SA Level II
Experience has a direct, positive effect on SA Level I
Experience has a direct, positive effect on SA Level II
(supplementary hypotheses) Moderating effects of age on VR psychological experience factors
H4: Age has a direct, positive effect on interactivity
H5: Age has a direct, negative effect on telepresence
H6: Age does not have a direct effect on immersion



Note: The dashed line and +/- icon represents an expected non-significant relationship between age and immersion

Figure 3. Proposed model of psychological and pilot factors influencing pilot SA in virtual reality.

4. Method

All testing took place in the Advanced Cognitive Engineering Laboratory at Carleton University, with ethics approval granted by Carleton University Ethics Approval Board-B (see Appendix A). The research discussed in this thesis is a subset of a study exploring the development of cognitive testing tools for general aviation pilots. The larger study involved participants completing two test sessions, the first in a traditional flight simulator, and the second in a VR flight simulator. As a result of this arrangement, completion of introduction and consent procedures, as well as collection of demographics information was carried out during session one. However, only methodological considerations and findings relevant to the second session are covered in this thesis.

4.1. Participants

Data was collected from a total of 47 participants. Fifty-one individuals completed session one testing and submitted demographics information, though four were unable to return for the second session; as a result, this data is omitted from the following discussion. The majority of participants were recruited through flyers posted at Ottawa-region flight clubs in Ontario, Canada. In addition, some participants were contacted as the result of having completed earlier ACE lab aviation studies and providing consent to be added to a mailing list for laboratory research. Eligibility criteria for participation in the study included individuals who were in possession of a valid pilot's license or permit and medical certificate, and had flown as pilot in command within the past two years. A summary of participant demographic information is provided below in Table 1. Participants were between 17 and 71 years old (M = 47.13, SD = 17.42), had logged between 2 and 12,000 total flight hours (M = 1384.85, SD = 2684.51), had served between 0 and 582 hours as pilot in command within the last two years (i.e., recent hours) (*M* = 14.77, *SD* = 14.34), and had been licensed between 1 and 70 years (*M* = 50.83, *SD* = 100.12). Pilot licensure was categorized into six levels, ranging from 1- student pilot, to 6- airline, air transport, or military pilot. The most common license/rating was "4 - visual flight rules with additional ratings" (see Table 2 Notes for the full list of ratings). Food and beverages were provided in lieu of financial compensation.

Table 2.

Participant Characteristics

	Age	Total Hours Flown	Recent Hours	Total Years Licensed	License Level
Minimum	17	2	0	1	1
Maximum	70	12,000	582	70	6
Mean	47.13	1384.85	50.83	14.77	4.09
SD	17.42	2684.51	100.12	14.34	1.38

Notes. License level was based on a six-point scale, where 1 - student permit, 2 – recreational permit, 3 – visual flight rules (no additional ratings), 4 – visual flight rules with additional ratings, 5 – instrument rated, commercial, and instructors; and 6 – airline transport.

4.2. Materials

During the flight exercise, participants wore an Oculus Rift headset and flew in a simulated Cessna 172 aircraft, which was rendered using Lockheed Martin's Prepar3D software. In place of a traditional simulator cockpit, the ACE lab engineering team created a flight control unit (Figure 4) using an adjustable seat, a set of rudder pedals, and a yoke, throttle, and flaps, which were installed into a wooden panel and set atop a desk. Only the primary flight instruments ("the 6-pack") and these engine controls were responsive in the virtual aircraft. Participants were also provided with a laboratory computer to complete all session questionnaires, which were generated using Google

Forms. A second, separate component of this session involved performance and work load testing using NASA's Multi-Attribute Task Battery II software.

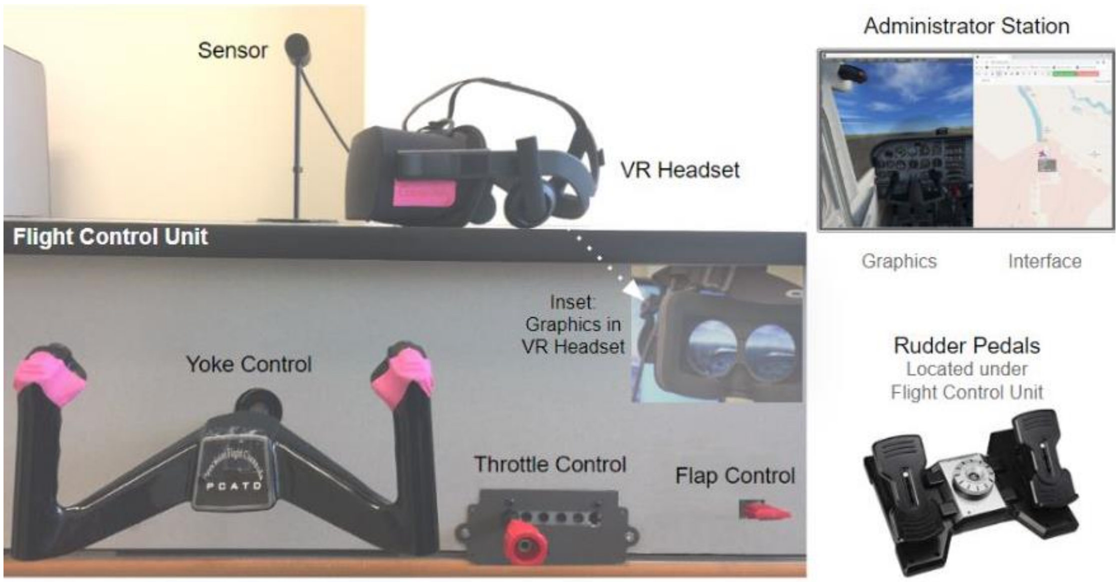


Figure 4. (R-L, top-bottom) Oculus Rift VR headset, sensor, and prototype flight control unit, including yoke, throttle, and flaps; administrator view of the simulated environment and geographical tracking of flight; and rudder pedals accompanying flight control.

4.3. Procedure

4.3.1. Introduction and Demographics Questionnaire

During the first session, all participants were given an introductory presentation to communicate the purpose of the study and provided an overview of experimental tasks and equipment. Participants were also given a paper copy of the study consent form (see Appendix B) and were asked to provide written confirmation of their intention to participate in both sessions. Researchers verbally reiterated the right of voluntary participation and provided an opportunity for participants to express any questions or concerns. Participants were also reminded of the small potentiality for experiencing negative simulator sensations (e.g., queasiness, dizziness, etc.), and were asked to alert researchers at the onset of any such symptoms. Following the presentation, participants were asked to complete a questionnaire, which was designed to collect general background information, such as age, as well as information pertaining to each participant's flight experience, including total number of flight hours, years licensed, and highest level of licensing at the time of the study.

4.3.2. Training

Participants were asked to complete a preliminary practice flight at the simulated Rockcliffe Aerodrome (CYRO) in Ottawa, Ontario, Canada. The practice flight consisted of two left-hand circuits, completed without any additional task requirements; this exercise was intended to help participants become accustomed to the VR environment and to any sensitivities created by the prototype controls. Time was also taken to ensure the comfort of the headset and to make any adjustments needed to improve visual clarity.

4.3.3. Flight Mission

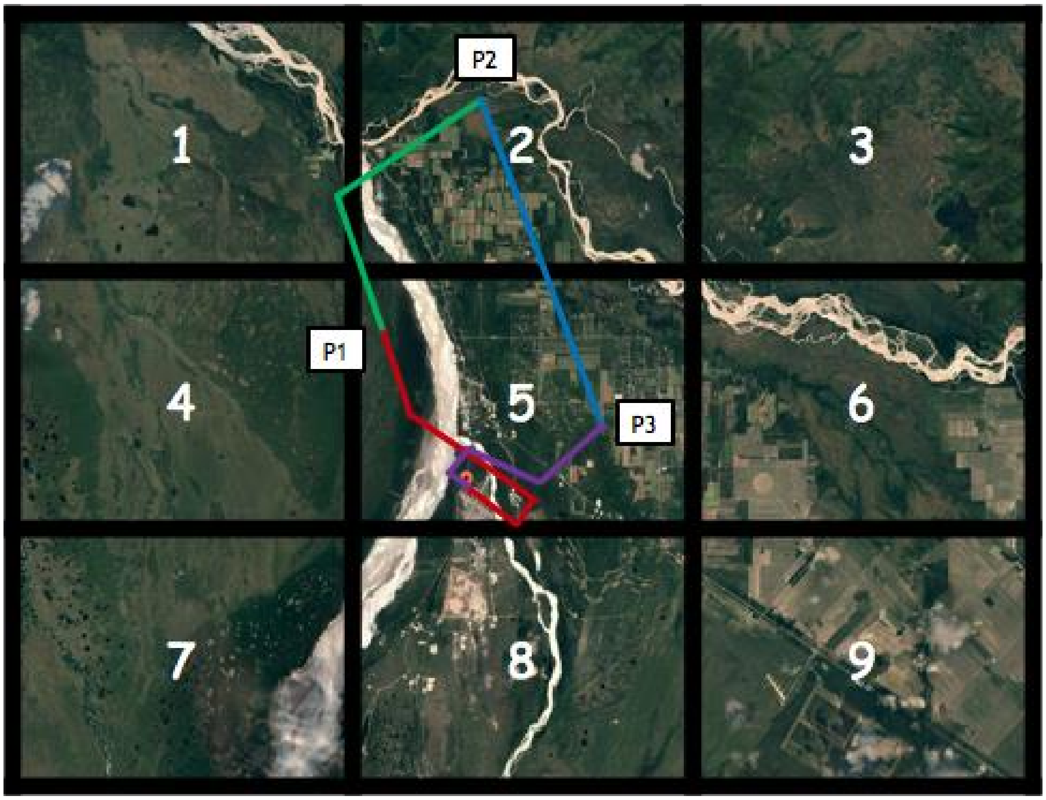


Figure 5. Topographical view of pilot flight path partitioned into zones, where P1-P3 represent planned SAGAT freezes.

The flight mission lasted ~35-45 minutes, and consisted of a four-leg route, which was completed in environmental conditions clear enough to permit the use of visual flight rules. Prior to the beginning of the exercise, participants were provided a general flight plan and notified that they would be given specific directions pertaining to airspeed, heading, and altitude over the duration of the flight. The exercise was designed in this manner to avoid participants flying off course, and to ensure that participants were not advantaged or disadvantaged in completing the experimental tasks. Individuals were also shown a map of the location of their flight and relevant airports, which was partitioned into nine zones (see Figure 5). A similar map was represented as an instrument within the virtual Cessna cockpit, and participants were asked to mentally track their zone placement throughout the flight. Participants were notified that though they would be hearing radio calls coming over the headset, the airfield in the virtual environment was uncontrolled, and other aircraft were not visually represented in the simulation. Participants were advised to read back all instructions they were given to ensure that the information was accurately communicated.

4.3.4. Situation Awareness Task

An SA-testing protocol was developed in accordance with recommendations made by Endsley (1995c) for the Situation Awareness Global Assessment Technique (SAGAT). In particular, freezes occurred at random intervals, with system displays being removed from the participant's field of view, and questions being administered immediately. SA freezes occurred at three different points during the flight (see markers P1-P3 above in Figure 5). Using controls on the Prepar3D interface, the simulation was paused, and the visual display of the Oculus Rift was temporary turned off. During this time, participants remained seated with the headset on, while researchers administered a set of questions related to pilot SA called the CanFly queries (Table 3). To avoid any distraction or disorientation caused by the freeze, researchers provided participants with a quick verbal warning that a pause was occurring and that their visual display would appear completely black.

Participants were administered a total of 24 questions, which were delivered across three SA freezes. Questions tested participant SA at three levels according to Endsley's model (1988), with 16 queries focusing on Level I SA (e.g., "State the call sign of the aircraft who most recently reported their position"), 6 on Level II (e.g., "Your current position is in what zone?") and 2 on Level III (e.g., "What aircraft was of most relevance to you? State the type and location."). The larger proportion of questions at the first level of SA aligned with findings from Endsley and Jones (2012), which report that ~75% of SA errors can be attributed to issues in Level I SA, while only 19% can be categorized at Level II, and 6% at Level III. All participants received queries in the same order and experienced SA freezes at approximately the same time frame within the flight. Following data collection, individual scores were calculated for each CanFly query item. Table 3 indicates sample queries for each level of SA.

Table 3.

Sample CanFly Queries and Level of SA

Question	SA Type
State the call sign of a particular aircraft	1
State the last reported position of a particular aircraft	1
Indicate your current position	2
State your current airspeed	1
State the intended actions of a particular aircraft	1
What was the reported wind direction	1
What aircraft was of most relevance to you... type and location	2 and 3

4.3.5. VR Psychological Experiences Questionnaire

VR psychological experiences are typically evaluated using subjective measures, with post-test questionnaires representing the most common method adopted by researchers, especially in the context of telepresence research (van Baren and IJsselsteijn, 2004). Sheridan (1992) suggests that since such experiences are "mental manifestations" they ought to be evaluated through subjective methods. In the case of users' psychological experiences of VR environments, post-test questionnaires generally offer a high level of validity (i.e. they appear to measure the intended concept), are simple to administer, and relatively straight-forward to analyze (van Baren and IJsselsteijn, 2004).

In the present study, participants were administered a 12-item post-test questionnaire (shown below in Table 4), which was adapted from the work of Mütterlein (2018). These items assessed user's subjective psychological experiences within the VR environment across 3 constructs: telepresence (e.g. "The content of the Oculus Rift seemed to be somewhere I visited rather than something I saw"), interactivity (e.g., "I had the feeling that I could influence the virtual world of the Oculus Rift"), and flow-based immersion, the latter of considered immersion across two dimensions, absorption by activity (e.g., "I didn't notice time passing") and fluency of performance (e.g., "My thoughts and movements felt effortless").

Table 4.

VR Psychological Experiences Questionnaire

No.	Item
Telepresence	
1	The Oculus Rift created a new world for me, and this new world suddenly disappeared when the exercise ended. Strongly disagree 1 2 3 4 5 6 7 Strongly agree
2	When I removed the Oculus Rift, I felt as if I returned to the "real world" after a journey. Strongly disagree 1 2 3 4 5 6 7 Strongly agree
3	I forgot about my immediate surroundings when I was using the Oculus Rift. Strongly disagree 1 2 3 4 5 6 7 Strongly agree
4	The content of the Oculus Rift seemed to be "somewhere I visited" rather than "something I saw". Strongly disagree 1 2 3 4 5 6 7 Strongly agree
Immersion: Fluency of Performance	
5	I had no difficulty concentrating. Strongly disagree 1 2 3 4 5 6 7 Strongly agree
6	My mind was free to focus on flying Strongly disagree 1 2 3 4 5 6 7 Strongly agree
7	My thoughts and movements felt effortless. Strongly disagree 1 2 3 4 5 6 7 Strongly agree
Immersion: Absorption	
8	I didn't notice time passing. Strongly disagree 1 2 3 4 5 6 7 Strongly agree
9	I was totally absorbed in what I was doing. Strongly disagree 1 2 3 4 5 6 7 Strongly agree
Interactivity	
10	The Oculus Rift content allowed me to interact with the virtual world. Strongly disagree 1 2 3 4 5 6 7 Strongly agree

11	I had the feeling that I could influence the virtual world of the Oculus Rift. Strongly disagree 1 2 3 4 5 6 7 Strongly agree
12	The Oculus Rift content was interactive. Strongly disagree 1 2 3 4 5 6 7 Strongly agree

4.3.6. Debriefing

Following the completion of the experimental exercises, participants were given time to rest and pose any questions they might have to the attending researcher. A form with all pertinent study and contact information was also provided to participants at the end of the session.

5. Results

Path analyses using partial least squares modeling were conducted to investigate relationships between three attributes of pilots' psychological experiences during VR exposure, telepresence, interactivity, and immersion (evaluated across two dimensions, fluency of performance and absorption by activity), as well as relationships to participants' Level I and II SA. Effects of age and experience on SA were also examined, as well as the effects of age on participants' VR experiences. Path analysis using partial least squares modeling was chosen for its suitability in handling datasets derived from a small population where mediating and moderating effects are of interest (Hair et al., 2014). Results were analyzed using WarpPLUS v. 6.0 (Kock, 2018). It should be noted that the reported beta values for the direct and indirect relationships discussed in this paper do not necessarily represent linear relationships. The underlying algorithms in the modeling software search for the relationship that best fits a pair of latent variables, which may be non-linear.

5.1. Validating the Measurement Models

The first step in any the model-building process is the construction of valid measurement models. As will be shown below, all factors in the final model (see Figure 8) showed acceptable quality in terms of factor loadings or weights and demonstrated the expected relationships with other factors.

5.1.1. VR Psychological Experience Measurement Model Validation

Regarding the proposed VR psychological experience constructs, responses to the 12-item questionnaire were first analyzed in SPSS using partial using principal component analysis (PCA) with oblique (direct oblimin) rotation. PCA was chosen in this context instead of confirmatory factor analysis, as the language in the questionnaire items was changed from that of the original instrument used in Mütterlein (2018) to better fit the current study. Moreover, the method for the present research and population tested was different from that of Mütterlein (2018). Oblique rotation was chosen over orthogonal rotation, as the VR experience factors were expected to be correlated. All items correlated at .30 or greater with at least one other item, suggesting reasonable factorability. Correlations are shown below in Table 5.

Table 5.

Correlations Among VR Psychological Experience Questionnaire Items

No.	Description	2	3	4	5	6	7	8	9	10	11	12
1	The Oculus Rift created a new world for me, and this new world suddenly disappeared when the exercise ended.	.66**	.40**	.51**	.28*	.15	.21	.04	.28*	.25*	.29*	.09
2	When I removed the Oculus Rift, I felt as if I returned to the "real world" after a journey.	-	.32*	.38**	.29*	.04	.21	.24	.28*	.17	.38**	.08
3	I forgot about my immediate surroundings when I was using the Oculus Rift.		-	.37**	.25*	.07	.10	.38**	.43**	.51**	.47**	.41**
4	The content of the Oculus Rift seemed to be "somewhere I visited" rather than "something I saw".			-	.34**	.03	.06	.10	.13	.25*	.19	-.12
5	I had no difficulty concentrating.				-	.58**	.45**	.08	.36**	.06	.30*	.01
6	My mind was free to focus on flying.					-	.62**	.05	.37**	.13	.26*	.17
7	My thoughts and movements felt effortless.						-	.21	.44**	.09	.42**	.17
8	I didn't notice time passing.							-	.46**	.36**	.36**	.39**
9	I was totally absorbed in what I was doing.								-	.55**	.61**	.57**
10	The Oculus Rift content allowed me to interact with the virtual world.									-	.61**	.69**
11	I had the feeling that I could influence the virtual world of the Oculus Rift.										-	.48**
12	The Oculus Rift content was interactive.											-

*.Correlation is significant at the .05 level (2-tailed), **.Correlation is significant at the .01 level (2-tailed)

In addition, the Kaiser-Meyer-Olkin measure verified the sampling adequacy for the analysis, KMO = .74, which is above the commonly acknowledged threshold of .60. When considering the anti-image correlation matrix (another measure of sampling adequacy), diagonal values for all items exceeded the acceptable level of .50. Three factors were identified with Eigenvalues over 1.0, which together explain 66.66% of the variance in users' VR psychological experience assessments. Table 6 shows the rotated factor loadings for each item. Factor one has been identified as "interactivity", factor two as "telepresence", and factor three as "immersion–fluency".

Table 6.

Rotated Factor Loadings for Each VR Psychological Experience Questionnaire Item

No.	Item	Mean	SD	F1	F2	F3
Telepresence						
1	The Oculus Rift created a new world for me, and this new world suddenly disappeared when the exercise ended.	4.47	1.73	.02	**.83**	.07
2	When I removed the Oculus Rift, I felt as if I returned to the "real world" after a journey.	4.51	1.68	.08	**.75**	.04
3	I forgot about my immediate surroundings when I was using the Oculus Rift.	5.00	1.73	**.57**	.43	-.11
4	The content of the Oculus Rift seemed to be "somewhere I visited" rather than "something I saw".	4.12	1.74	-.07	**.82**	-.04
Immersion: Fluency of performance						
5	I had no difficulty concentrating.	5.51	1.52	-.11	.31	**.73**
6	My mind was free to focus on flying.	5.89	1.29	.02	-.13	**.90**
7	My thoughts and movements felt effortless.	5.40	1.53	.10	-.04	**.82**
Immersion: Absorption by activity						
8	I didn't notice time passing.	5.23	1.66	**.64**	-.01	-.02
9	I was totally absorbed in what I was doing.	5.98	1.00	**.68**	.03	.36
Interactivity						
10	The Oculus Rift content allowed me to interact with the virtual world.	5.85	1.08	**.85**	.09	-.12
11	I had the feeling that I could influence the virtual world of the Oculus Rift.	4.62	1.80	**.65**	.16	.25
12	The Oculus Rift content was interactive.	5.60	1.26	**.90**	-.25	-.01
Eigenvalues				**4.35**	**1.91**	**1.74**
% of variance				**36.25**	**15.92**	**14.50**

Chronbach's Alpha	**.82**	**.76**	**.78**

Notes: PCA converged in 7 iterations; factor loadings > .50 shown in bold

Three factors were created to represent VR psychological experiences: interactivity, telepresence, and immersion–fluency. Items 3 and 9-12 were used as indictors for interactivity; items 1, 2, and 4 as indicators for telepresence; and items 5-7 as indicators for immersion–fluency. Item 8 from the VR psychological experiences questionnaire (relating to perceived passage of time) was eliminated during the measurement model validation phase, as its addition lowered the composite reliability coefficient and average variance extracted (AVE) for the interactivity factor.

Next, a simple model including only the three VR experience factors was tested. As shown in Figure 6, interactivity was observed to have direct, positive effects on both telepresence, $\beta = .49$, $p < .01$, as well as immersion–fluency, $\beta = .41$, $p < .01$. Moreover, telepresence had a direct, positive effect on immersion–fluency, $\beta = .28$, $p < .01$. Telepresence was also shown to have a significant mediating effect on the relationship between interactivity and immersion–fluency, $\beta = .14$, $p =.02$.

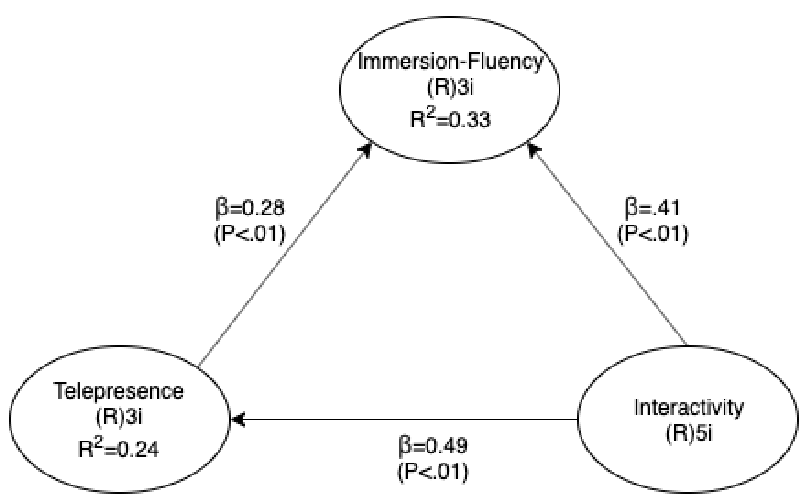


Figure 6. Model of relationships among VR psychological experiences.

5.1.2. Experience Measurement Model Validation

Experience was a formative variable, comprising indicators made up of pilots' certification level, total years licensed, total hours flown, and hours flown as pilot in command within the last two years. In the case of formative variables, indictors are not expected to be correlated with one another. As shown in Table 7, the weights for each indicator were significant, and were thus retained in the final measurement model.

5.1.3. Situation Awareness Measurement Model Validation

Two reflective latent variables were initially created for Level I and II SA, which contained all items from the CanFly queries. Increasingly granular SA constructs were created on the basis of the similarity of question type, and indicators were removed if they resulted in factor loadings below .50, or if they reduced the AVE score of the latent variable when included. The Level I SA variable involves processing of auditory information related to the activity of neighboring aircraft, comprising indicators made up of two items representing the relative position of other aircraft. The Level II SA variable involves comprehension of the location of one's ownship relative to neighbouring aircraft and was comprised of two indicators reflecting ownship position and the details of other aircraft most relevant to the trajectory of the participant. An examination of the latent variable coefficients revealed composite reliability coefficients of .67 and .78, and AVE scores of .50 and .64, confirming the viability of both SA Level I and II variables, respectively. Other similar items with less than ideal loadings were removed from the measurement model in order to present the strongest SA variables possible.

As shown below in Figure 7, relationships were evaluated among the two SA factors and with age and experience to confirm ecologically validity of the factors.

Previous work has shown that SA Level I should directly predict SA Level II; moreover, while pilot age tends to negatively impact SA at both the first and second level, experience should positively affect SA at both levels (Van Benthem et al., 2011). SA Level I was observed to have a direct positive effect on SA Level II, $\beta = .32$, $p < .01$. Age was shown to have a direct, negative impact on both Level I SA, $\beta = -.22$, $p = .01$, as well as Level II SA, $\beta = -.39$, $p < .01$. Finally, as expected, experience had a significant, positive effect on SA Level I, $\beta = .27$, $p < .01$, as well as SA Level II, $\beta = .23$, $p < .01$.

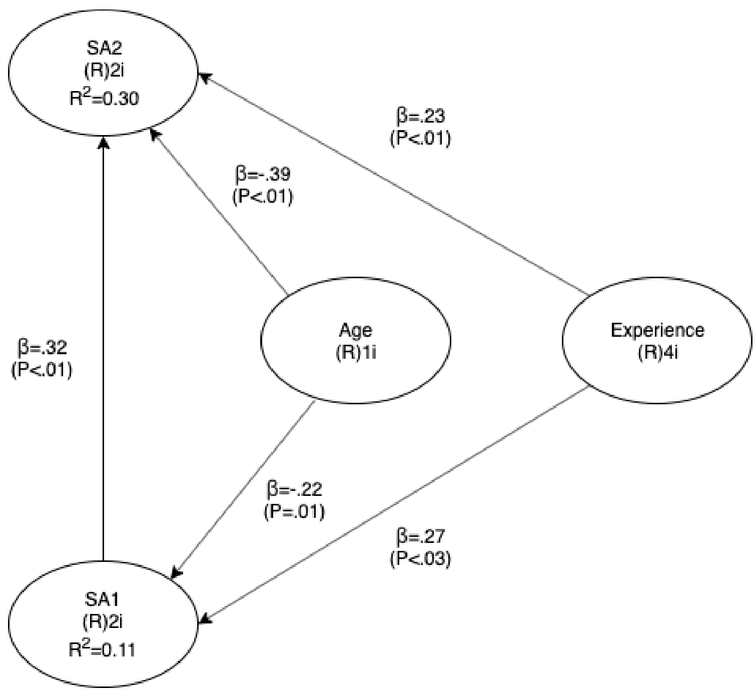


Figure 7. Model of relationships among SA Level I, SA Level II, age, and experience.

5.1.4. Measurement Model Summary

In summary, six reflective latent variables and one formative variable were integrated into the model. All reflective latent variables had composite reliability coefficients above .70 and resulted in AVE values greater than or equal to the recommended .50 threshold (Kock, 2018). Composite reliability coefficients and AVE are not reported for formative indicators. The indicators for the reflective variables used in the subsequent analysis met the acceptability criteria outlined by Kock (2018), which include factor loadings above .5 and *p*-values below .05. Regarding the indicators for the formative experience construct, indicator weights and *p*-values are reported, where weights with *p*-values below .05 were considered valid. See Table 7 for a complete list of indicator acceptability criteria values, and Table 8 for a list of the AVE values and composite reliability coefficients for all reflective variables in the model.

Table 7. *Indicator Acceptability Criteria Values*

Variable	Indicator description	Indicator weights	Factor loadings	*p*-value
Experience*	Total hours flown	.32	-	<.001
Experience*	Recent hours	.32	-	<.001
Experience*	Total years licensed	.32	-	<.001
Experience*	License level	.32	-	<.001
Telepresence	The Oculus Rift created a new world for me, and this new world suddenly disappeared when the exercise ended.	-	.89	<.001
Telepresence	When I removed the Oculus Rift, I felt as if I returned to the "real world" after a journey.	-	.83	<.001
Telepresence	The content of the Oculus Rift seemed to be "somewhere I visited" rather than "something I saw".	-	.74	.006
Immersion-fl.	I had no difficulty concentrating.	-	.80	<.001
Immersion-fl.	My mind was free to focus on flying.	-	.89	<.001
Immersion-fl.	My thoughts and movements felt effortless.	-	.82	<.001

Interactivity	I forgot about my immediate surroundings when I was using the Oculus Rift.	-	.69	<.001
Interactivity	I was totally absorbed in what I was doing.	-	.80	<.001
Interactivity	The Oculus Rift content allowed me to interact with the virtual world.	-	.86	<.001
Interactivity	I had the feeling that I could influence the virtual world of the Oculus Rift.	-	.80	<.001
Interactivity	The Oculus Rift content was interactive.	-	.80	<.001
SA1	State the last reported position of Papa Hotel Victor	-	.80	<.001
SA1	State the last reported altitude of Gulf Hotel India	-	.80	<.001
SA2	Your current position is in what zone	-	.79	<.001
SA2	What aircraft was of most relevance to you in this last leg of flight... state the type and location	-	.79	<.001

Note: Indicator weights are reported for formative variables

Table 8.

Average Variance Extracted and Composite Reliability Coefficients for Reflective Models

Variable	Average variance extracted	Composite reliability coefficient
Telepresence	.68	.86
Immersion–fluency	.70	.88
Interactivity	.63	.89
SA1	.50	.67
SA2	.64	.78

5.2. Evaluating the Full Model of Psychological and Pilot Factors on Situation Awareness in Virtual Reality

5.2.1. Model Quality Assessment

In order to test the study's main hypotheses, a combined model of all psychological and pilot factors was developed. Model data were ranked to mitigate against idiosyncratic effects of outliers. During the initial structural modelling process, the fit and quality of the model was evaluated using the 10 global indices recommended by Kock (2018): average path coefficient (APC), average R-squared (ARS), average adjusted R-squared (AARS), average block variance inflation factor (AVIF), average full collinearity variance inflation factor (AFVIF), Tenenhaus GoF (GoF), Simpson's paradox ratio (SPR), R-squared contribution ratio (RSCR), statistical suppression ratio (SSR), and non-bivariate causality direction ratio (NLBCDR). For APC, ARS, and AARS, p-values are assessed, where $p \leq .05$ is considered the threshold for acceptance. APC and ARS will only increase together if new latent variables added to the model enhance its overall predictive value. AARS corrects for erroneous increases in R-squared coefficients stemming from low-value variables. High AVIF and AFVIF values indicate a potential overlap in the meaning of model variables. It is recommended that both values be equal to or lower than 3.3. GoF is a measure of a model's explanatory power. Wetzels et al. (2009) provides the following threshold levels: values equal to or greater than .1 and less than .25 are small, values equal to or greater than .25 and less than .36 are medium, and values equal to or greater than .36 are large. Any values below .1 indicate that the explanatory power of the model is too low to be viable. The SPR represents the extent to which paths in the model are free of instances of Simpson's paradox; the latter occurs when a path coefficient and correlation related to two variables have different signs. A value of 1.00 is considered ideal. Path-correlation signs were therefore also investigated, and a negative sign was identified for the path between age and telepresence, suggesting an incidence of Simpson's paradox. These conditions suggest that the path between age and telepresence is implausible or reversed, so it was removed from the model. RSCR is evaluated alongside the SPR and measures the extent to which a model is free of negative R-squared contributions, which suggest that the percentage of variance explained is being reduced by a predictor. A value of 1.00 is considered ideal for RSCR. SSR assesses the extent to which the model is free of instances of statistical suppression, which like SPR can indicate causality issues. An SPR value greater than or equal to .70 is considered acceptable. NLBCDR evaluates the extent to which "bivariate non-linear coefficients of association" support the direction of the causal links in the model. NLBCDR values greater than or equal to .70 are considered desirable. Following the elimination of the age and telepresence path, all global fit and quality indices were well within satisfactory ranges. A detailed summary of these criteria is provided below in Table 9.

Table 9.

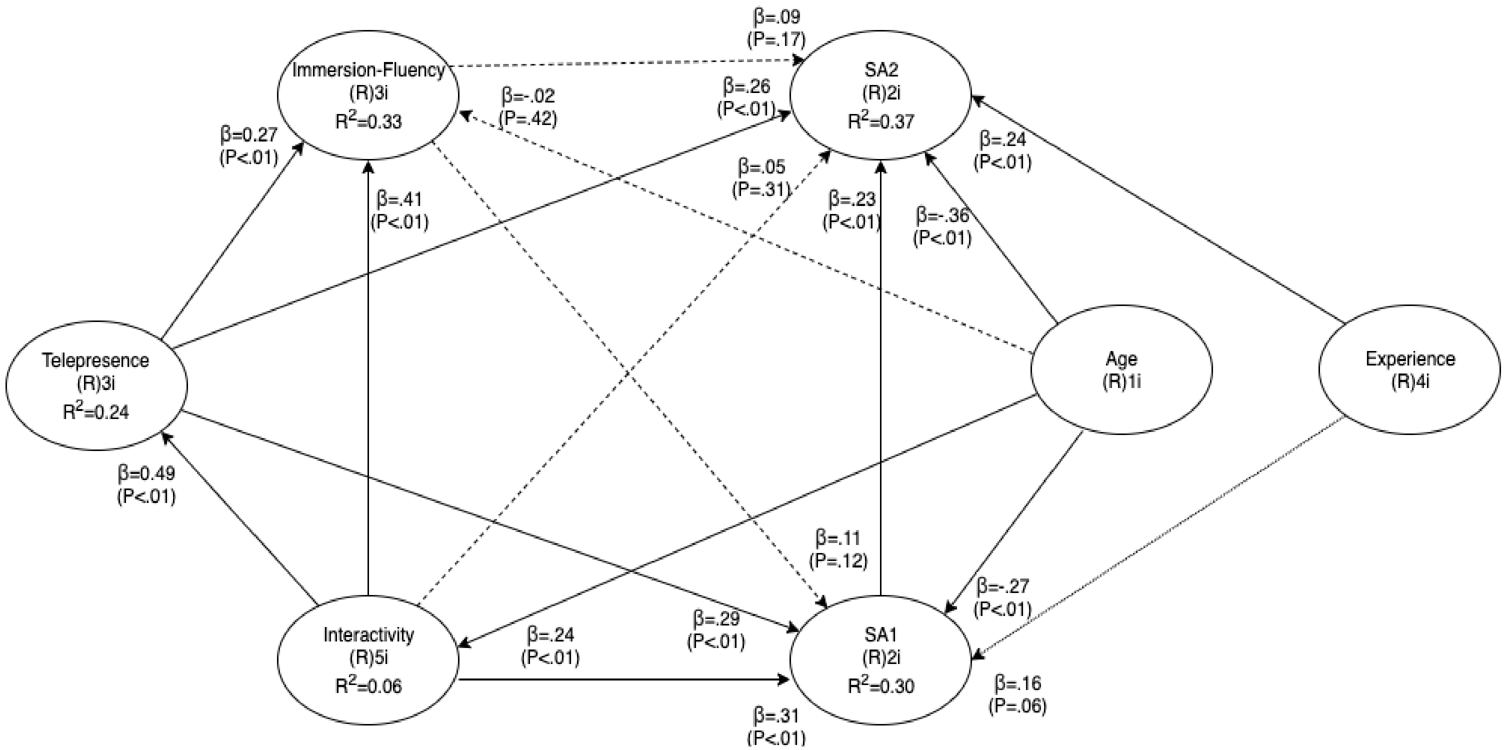
Global Fit and Quality Indices

Index	Value
Average path coefficient (APC)*	.24
Average R-squared (ARS)*	.26
Average adjusted R-squared (AARS)*	.21
Average block VIF	1.18
Average full collinearity VIF	1.5
Tenenhaus GoF	.43
Simpson's paradox ratio	1.00
R-squared contribution ratio	1.00
Statistical suppression ratio	.88
Nonlinear bivariate causality direction ratio	.91

Note: p-values are also reported for APC, ARS, and AARS; in all cases, $p < .01$

5.2.2. Structural Model Evaluation

The effects of telepresence, interactivity, and immersion–fluency on pilot SA at the first and second level were evaluated. Interactivity was shown to have a direct, positive impact on pilots' SA Level I, $\beta = .31$, $p < .01$, validating hypothesis 1a. However, no direct relationship was observed between interactivity and SA Level II; hypothesis 1b was therefore not supported. Indirect effects of telepresence and immersion–fluency also explain a significant portion of the relationship between interactivity and SA Level I, $\beta = .19$, $p = .03$. In addition, indirect effects of telepresence, immersion–fluency, and SA Level I explain a significant portion of the relationship between interactivity and SA Level II, $\beta = .24$, $p = .01$. Telepresence had a direct positive impact on pilots' SA Level I, $\beta = .29$, $p < .01$, and SA Level II, $\beta = .26$, $p < .01$, which provides support for hypotheses 2a and 2b. No notable effects were observed between immersion–fluency and SA Level I or II. Therefore, neither of hypotheses 3a and 3b were supported by the present study. In addition, when the VR psychological experience variables were integrated into the combined model, the relationship between experience and SA Level I fell below the threshold of significance, SA Level I, $\beta = .16$, $p = .06$. All direct relationships are illustrated below in Figure 8. In addition, values for all direct, indirect, and total relationships relevant to the model are listed in Tables 10 through 12. Table 13 also shows effect sizes for total effects.



Note: Solid lines represent significant relationships, dashed lines represent non-significant relationships, and the dotted line (between experience and SA Level I) represents a weak, though meaningful relationship

Figure 8. Model of psychological and pilot factors influencing pilot SA in virtual reality.

Table 10.

Direct Beta Values and p-values for All Latent Variables in the Model

	Telepresence	Immersion–fluency	Interactivity	SA Level 1	SA Level 2	Age	Experience
Telepresence	-	-	.49 (<.01)	-	-	-	-
Immersion–fluency	.27 (<.01)	-	.41 (<.01)	-	-	-.02 (.42)	-
Interactivity	-	-	-	-	-	.24 (<.01)	-
SA Level 1	.29 (<.01)	.11 (.12)	.31 (<.01)	-	-	-.27 (<.01)	.16 (.06)
SA Level 2	.26 (<.01)	.09 (.17)	.05 (.31)	.23 (<.01)	-	-.36 (<.01)	.24 (<.01)
Age	-	-	-	-	-	-	-
Experience	-	-	-	-	-	-	-

Note: *p*-values are displayed in brackets

Table 11.

Indirect Beta Values, Number of Paths, and p-values for Paths with Two Segments

	Telepresence	Immersion–fluency	Interactivity	SA Level 1	SA Level 2	Age	Experience
Telepresence	-	-	-	-	-	.12,1, (.05)	-
Immersion–fluency	-	-	.13, 1, (.03)	-	-	.10, 1, (.08)	-
Interactivity	-	-	-	-	-	-	-
SA Level 1	.03, 1, (.33)	-	.19, 2, (.03)	-	-	.07, 2, (.23)	-
SA Level 2	.09, 2, (.17)	.03, 1, (.35)	.24, 3, (.01)	-	-	-.08, 3, (.22)	.04, 1, (.30)
Age	-	-	-	-	-	-	-
Experience	-	-	-	-	-	-	-

Note: *p*-values are displayed in brackets

Table 12.

Total Beta Values, Number of Paths, and p-values for All Latent Variables in the Model

	Telepresence	Immersion–fluency	Interactivity	SA Level 1	SA Level 2	Age	Experience
Telepresence	-	-	.49, 1, (<.001)	-	-	.12, 1,	-
Immersion–fluency	.27, 1, (.003)	-	.55, 2, (<.001)	-	-	.11, 3, (.24)	-
Interactivity	-	-	-	-	-	.24, 1, (.003)	-
SA Level 1	.31, 2, (.001)	.11, 1, (.06)	.51, 4, (<.001)	-	-	-.15, 6, (.02)	.16, 1, (.03)
SA Level 2	.36, 4, (<.001)	.12, 2, (.25)	.25, 8, (.10)	.23, 1, (.01)	-	-.37, 12, (<.001)	.28, 2, (.35)
Age	-	-	-	-	-	-	-
Experience	-	-	-	-	-	-	-

Note: *p*-values are displayed in brackets

Table 13.

Effect Sizes for Total Effects of All Latent Variables in the Model

	Telepresence	Immersion–fluency	Interactivity	SA Level 1	SA Level 2	Age	Experience
Telepresence	-	-	.24	-	-	.06	-
Immersion–fluency	.12	-	.28	-	-	.02	-
Interactivity	-	-	-	-	-	.06	-
SA Level 1	.10	.04	.13	-	-	.03	.04
SA Level 2	.13	.02	.03	.08	-	.15	.04
Age	-	-	-	-	-	-	-
Experience	-	-	-	-	-	-	-

Note: According to guidelines by Cohen (1988), *f*≥.02, *f*≥.15, and *f*≥.35 represent small, medium, and large effect sizes, respectively.

5.3. Supplementary Hypotheses: Evaluating the Effects of Age on Virtual Reality Experience Factors

To test study's supplementary hypotheses, the effects of age on telepresence, interactivity, and immersion–fluency were explored. These results also helped to identify any peripheral moderating influences on the relationships between the VR psychological experience factors and SA. As was discussed above, the path between age and telepresence was removed from the model due to an incidence of Simpson's paradox; therefore, hypothesis 4 was not supported by the present study. However, age was shown to have a significant indirect effect on telepresence through interactivity, $\beta = .12$, $p = .05$. In contrast, age had a direct positive effect on interactivity, $\beta = .24$, $p < .01$, supporting hypothesis 5. These results suggest that older pilots tended to perceive a greater sense of control over the system. Age did not have a significant direct effect on immersion–fluency, confirming hypothesis 6. All direct relationships discussed in this section are illustrated above in Figure 8.

5.4. Summary of Hypotheses

Table 14 outlines a summary of outcomes for all relationships tested in the combined model.

Table 14.

Summary of Relationships Supported and Not Supported by the Analysis

Relationships Tested	**Outcome**
Relationships between VR psychological experience factors and SA	
H1a: Interactivity has a direct, positive effect on SA Level I	Supported
H1b: Interactivity has a direct, positive effect on SA Level II	*Not supported*
H2a: Telepresence has a direct, positive effect on SA Level I	Supported
H2b: Telepresence has a direct, positive effect on SA Level II	Supported
H3a: Immersion has a direct, positive effect on SA Level I	*Not supported*
H3b: Immersion has a direct, positive effect on SA Level II	*Not supported*
Steps:	
(one) Validation of VR psychological experience factors	
Interactivity has a direct, positive effect on telepresence	Supported
Interactivity has a direct, positive effect on immersion	Supported
Telepresence has a direct, positive effect on immersion	Supported
(two) Validation of SA factors	
SA Level I has a direct, positive effect on SA Level II	Supported
Age has a direct, negative effect on SA Level I	Supported
Age has a direct, negative effect on SA Level II	Supported
Experience has a direct, positive effect on SA Level I	Supported*
Experience has a direct, positive effect on SA Level II	Supported
(supplementary hypotheses) Moderating effects of age on VR psychological experience factors	
H4: Age has a direct, positive effect on interactivity	Supported
H5: Age has a direct, negative effect on telepresence	*Not supported*
H6: Age does not have a direct effect on immersion	Supported

Note: Though the direct relationship between experience and SA Level I fell below the level of significance when the relationship was tested in the combined model of factors, the total effect size (f = .04) remains above the threshold for viability.

6. Discussion

6.1. Relationships Among Virtual Reality Psychological Experience Factors

The first step in the analysis for the present study involved validation of three distinctive factors for users' psychological experiences in VR environments: telepresence, interactivity, and immersion–fluency. Relationships were also modeled among factors. Support was found for a similar version of the path analysis model proposed by Mütterlein (2018), with the same direct relationships found between telepresence, interactivity, and immersion–fluency. Interactivity was shown to have a direct positive effect on participants' perceived telepresence, while both telepresence and interactivity had direct positive effects on participants' levels of immersion–fluency. Moreover, the relationship between participants' perceived interactivity and immersion–fluency was mediated by their level of telepresence.

Mütterlein (2018) suggests that an individual must first feel as though they are present in a virtual environment to experience a sense of immersion during activities performed in the virtual environment. Findings from the present work also show that perceived telepresence is necessary for explaining, in part, the influence that interactivity has in enabling immersive experiences. Experiencing feedback from interactions and establishing a sense of control over content, as opposed to passively watching it, may be crucial for achieving a heightened sensation of being immersed in an activity (Mütterlein, 2018). There are a number of studies in the literature, which also support these explanations. For example, Hsu (2010) suggests that interactivity-type experiences, such as positive perceptions regarding ease of use, immediacy of system feedback, and match between skill and challenge, serve as antecedents to flow-based experiences (i.e., immersion). In addition, Hoffman and Novak (1996) assert that increased levels of interactivity and telepresence increase the “subjective intensity” of an individual’s flow state. Novak, Hoffman, and Yung (2000) later validated a model of flow, which identifies a direct positive effect of telepresence on flow. Following a study of users’ subjective experiences during VR game play, Weibel et al. (2008) also identified a strong positive relationship between telepresence and flow states.

Finally, the strength of an individual’s experience of interactivity seems to have a significant positive influence on the level of VR telepresence achieved by that individual. Similar relationships between interactivity and telepresence have also been reported by von der Pütten et al. (2012), who found that perceived interactivity predicted subjective experiences of telepresence in an augmented reality game. In addition, Siebert and Shafer (2017) found that the perceived naturalness of a controller (i.e., interactivity) had a significant positive effect on participants’ reported level of spatial presence during gameplay.

6.2. Relationships Among Situation Awareness Factors

6.2.1. Situation Awareness Levels I and II

The present study evaluated SA objectively using a SAGAT-style technique, which encompassed 24 queries administered over a series of three freezes. SA variables were developed selectively to achieve a set of viable indices with robust factor loadings and reliability coefficients. SA is known to be influenced by an individual’s goals and expectations (Endsley, 1997), and variation in these areas throughout each leg of flight may explain the relative lack of correlation among some indices for each Level of SA. However, two reflective latent variables were generated from these queries to develop the path analysis. The Level I SA variable comprises indices related to processing information about the activity of neighboring aircraft, while the Level II SA variable comprises indices related to comprehension of the location of one's ownship relative to neighbouring aircraft. As expected, results of the analysis showed a direct, positive relationship between these Level I and II SA variables. Furthermore, a number of direct and indirect effects of telepresence, interactivity, and immersion–fluency on SA Level I and II were observed.

6.2.2. Effects of Age and Experience

Given that the average age of the general aviation population is increasing, age is an important factor to consider when developing cognitive evaluation tools for GA pilots. Results from the present study demonstrated that age had direct negative effects on both SA Level I and II. These results were aligned with findings from the literature, which suggest that the information processing components that facilitate SA acquisition may be sensitive to age-related cognitive changes. At the first level of SA, age-related decline in resources related to attention and working memory may limit the amount of information that enters working memory, the efficiency of the retrieval processes, and the complexity of information that can be handled at any given moment (Bolstad & Hess, 2000). These changes may negatively influence an individual's ability to develop accurate mental models of their environment, and which support second- (comprehension) and third-level (projection) SA. In addition, age may negatively affect an individual's source memory, which supports the encoding of contextual information (Bolstad & Hess, 2000). Parsing informational context is essential for retention of information and is foundational for achieving higher levels of SA.

The effects of experience on the relationships between age and SA Level I and II were also examined. As was expected, experience was shown to have a direct, positive influence on SA at both levels. The present research provides support for the proposition that experience has a moderating effect on the relationship between age and SA. Experience may modify the extent to which age effects are observed in different contexts (Bolstad & Hess, 2000). There is evidence to suggest that the effect of age-related cognitive decline is reduced in situations in which an individual is highly experienced; extensive practice in a task environment supports the development of resources, including a broad knowledge base, automatization of operations, and mechanisms that help to compensate against the slowing of cognitive processes (Bolstad & Hess, 2000). For example, a 1984 study by Salthouse investigating age effects on typing speed found that though older adults were slower in terms of their key tapping rate and reaction time for key choice, their overall typing speed was comparable to that of their younger counterparts. The author suggests that the older adults tended to scan further ahead in the document they were typing sound, a behaviour which served to compensate for their limitations in processing speed. In the context of flight tasks, more experienced pilots may be better at anticipating certain cockpit conditions and be more proactive in reviewing their environment and instruments.

Hess and Slaughter (1990) assert that aging can be associated with increased reliance on long-term memory structures (e.g., mental models and schemata), which serve to compensate for decreasing efficiency in basic processing mechanisms. Mental models (complex structures that individuals use to model the behaviour of systems) and schemata (states of the mental model that organize categories of information and facilitate efficiencies in information processing) are thought to support SA by helping an individual focus on relevant information and form expectations regarding the task environment (Endsley & Jones, 2012). In general, schema use and activation appear to be minimally affected by age (Hess, 1990).

6.3. Investigating the Effects of Virtual Reality Psychological Experience Factors on Situation Awareness Level I and II

6.3.1. Telepresence, Interactivity, and Situation Awareness

The primary purpose of the present research was to model the direct and indirect effects of pilots' psychological experiences during VR exposure on their ability to achieve SA at the first and second level. Telepresence, the subjective experience of "being there" in a virtual environment, has long been treated as an indicator of the ecological validity of VR devices. Moreover, the phenomenon has been regarded as necessary for facilitating the transfer of knowledge acquired in virtual training conditions to real-world contexts (Mestre, 2006). Findings from the present study suggest that telepresence supports SA in VR environments at both the first and second level. Results showed direct, positive relationships between telepresence and SA Level I and between telepresence and SA Level II of comparable strengths. These results support similar findings from earlier discussed studies, which point to significant positive associations between telepresence and SA (Laptaned, 2006; Jung et al., 2008; and He et al., 2018).

In contrast, interactivity only directly affected SA at the first level (i.e. perception of elements within the environment), though it was shown to have a significant indirect effect on SA Level II through telepresence and immersion–fluency. Interactivity refers to a sense of agency, or perceived capacity to influence the form of content in a simulated environment (Mütterlein, 2018; Steur, 1992). Interactivity is likely to be influenced by attributes, such as the speed in which user input is represented in the environment, the range of actions available to the user, and the predictability of the responses generated by user inputs (Reed & Saleem, 2017). High levels of interactivity suggest that the user is focused on elements of the virtual environment and effectively filtering out external noise stimuli, both of which help to explain the positive effects of interactivity on telepresence and SA Level I. In contrast, interactivity may not have any impact on the activation of more complex cognitive structures, such as mental models, which enable an individual to achieve higher levels of SA. Endlsey and Jones (2012) underscore mental models as "key enablers" of complex SA processes (i.e., comprehension and projection).

Telepresence is thought to develop through the formation of more complex cognitive representations, which may provide a foundation for the achievement of higher-level SA. Schubert et al. (1998) and Schubert et al. (2001) maintain that experiences of telepresence emerge through the ongoing construction of a "spatial-functional" mental model. The authors suggest that two crucial cognitive processes facilitate the development of this mental model, namely the suppression of conflicting stimuli, and the continuous representation of bodily action in the virtual world. Our cognitive representations of an environment comprise "patterns of possible action", which are meshed with patterns of action based on memory. Meaning is then created through the interactions between our bodies and attributes of our environment, and the ongoing negotiation of those patterns of action. The importance of the representation of bodily action in this account also aligns with the strong, positive influence of perceived interactivity on telepresence experience observed in the present study.

However, it is also worth considering why it is the case that strong telepresence and interactivity experiences tend to support stronger SA rather than creating a source of resource competition? Endsley's model (1995a) depicts SA as a state of knowledge built up through a number of cognitive processes, including (but not limited to) working memory, which is also thought to be an important factor in supporting telepresence (Nunez, 2004). Similarly, Rawlinson, Lu, and Coleman (2012) suggest that individual differences in working memory affect an individuals' ability to achieve telepresence in a simulated environment. It may be tempting to assume that immersive VR systems that engage a greater range of sensory channels create an added demand on working memory. However, Nunez (2004) maintains that increased transmission of information does not necessarily consume additional working memory resources when that information can be easily grouped into meaningful chunks. The stimuli that enable an ongoing sensation of telepresence may exist within a more abstract conceptual complex, which contains powerful SA cues about the task context. When system latency occurs, the user is affected by an influx of stimuli from the external environment, which interrupts the normal chunking of information received from the simulated environment (Nunez, 2004). The discrepancy in stimuli flow in turn creates a tax on working memory and degrades perceived telepresence. Nunez (2004) proposes a two-phase processing model based on working memory to explain the occurrence of telepresence. In the first phase, perceptual stimuli are decoded and transformed into more abstract representations, such as schema or spatial-functional models. In the second phase, these representations form the basis of inferences about the virtual environment. The author suggests that telepresence is likely to arise as a by-product of this second phase of processing. The account from Nunez

(2004) is also compatible with those of Schubert et al. (1998) and Schubert et al. (2001), and provides another potential explanation for the direct, positive relationship between telepresence and Level II SA observed in the present study. Future research should explore relationships between working memory and VR psychological experiences, in addition to examining effects of working memory on the relationships between VR psychological experiences and SA.

6.3.2. Immersion–fluency and Situation Awareness

The results of the analysis showed there to be no significant relationship between immersion–fluency and SA at either level. The present research adopted a flow-based definition of immersion, whereby immersion is characterized as the subjective experience of feeling involved or absorbed in activities conducted in a virtual environment. Flow may be described as a psychological state of optimal experience, in which an individual is entirely absorbed or immersed in an activity (Csikszentmihalyi, 1990). The observed results were unexpected, given that flow-based experiences are thought to be components of meaningful learning in terms of depth of cognitive processing and performance (Shernoff and Csikszentmihalyi, 2009). Flow has been identified as a significant predictor of performance in the context of activities demanding a high level of attention and memory (Engeser, Rheinberg, Vollmeyer, & Bischoff, 2006; and Schüler, 2007). It was hypothesized that participants who achieved stronger flow states may remain more engaged and focused on stimuli relevant to the flight task, and in turn achieve greater accuracy when answering the SA queries.

Interestingly, the results of this study align with findings reported by Bian et al. (2018); despite the body of work supporting a positive association between flow experiences and performance in non-virtual environments, the relationship seems to fall away during assessment of tasks conducted using VR systems. Bian et al. (2018) suggest that the weak association between flow and performance may occur in VR training environments when interactive elements of the interface are overly fascinating, complex, or unintuitive, such that they draw the user's attention away from the primary task. In the present research context, it may be the case that the sensitivities of the flight control unit or attributes of the rendered environment created a distraction from the exercise. In addition, though pilots often heard information about neighbouring aircraft conveyed over their headset, no other aircraft were represented visually in the simulation, which removed a powerful visual cue that pilots often rely on. It was essential for the administration of the SA queries to avoid a mismatch between the visual behaviour of the aircraft and the information communicated over the pre-recorded radio calls. However, the Prepar3D interface available for the experimental exercise was unable to convey this information together in an integrated manner.

Moreover, given that the present work comes from a larger study exploring the development of VR cognitive evaluation tools for pilots, participants were also being asked to complete a prospective memory task, which contributed to the overall cognitive load placed on them during the exercise. Such a consideration may in turn be crucial for research design when evaluating effects of flow on SA, as it is thought to be difficult for individuals to achieve flow when they must perform at a state of cognitive overload (Ninaus, 2015). Flow is more likely to occur in conditions where skill is closely balanced with the demand of the task. Further evaluation of the effects of immersive states on pilot

SA during VR exposure should minimize added cognitive load created by the task environment.

6.4. Effects of Age on Virtual Reality Psychological Experience Factors

6.4.1. Age and Telepresence

A supplementary objective of the present study was to explore effects of age on participants' reported levels of telepresence, interactivity, and immersion–fluency. The present study was unable to provide support for a relationship between age and telepresence, as testing revealed an instance of Simpson's paradox, which signals causality issues. In the case of Simpson's paradox, a relationship may be present among sub-groups of a population but disappears or becomes reversed in the context of the aggregate population (Pearl, 1999). These results suggest that discussion of a direct relationship between age and telepresence is overlooking the influence of a third variable. Further research is needed to better understand the factors influencing telepresence experience.

6.4.2. Age and Interactivity

Results of the analysis found age to have a significant positive effect on participants' perceived interactivity. These findings were aligned with research from Bangay and Preston (1998), which noted that older individuals tended to report a stronger sense of control over the virtual environment. A potential explanation for this finding may be related to the expertise or general flight skill that older individuals brought to the exercise. Results from this study also suggest that absorption by activity, a dimension of flow, may be a component of interactivity. Flow-based experiences are thought to be more likely to occur in contexts where there is a close balance between an individual's skill and the demand of the activity. Studies by Ghani and Deshpande (1994) and Novak et al. (2000) provide support for a causal model of flow, which outlines direct relationships between skill and control and between control and flow. Further work should be undertaken to validate the conceptual relationship between skill/experience and perceived interactivity in VR environments.

6.4.3. Age and Immersion–fluency

Finally, as was initially hypothesized, no significant relationship was observed between age and immersion–fluency. These findings are aligned with those reported by Payne et al. (2011), which suggests that an individual's ability to achieve a flow state is not negatively compromised by age. The authors conclude that more cognitively demanding activities elicit higher levels of flow among individuals with greater fluid intelligence yet result in lower levels of flow for those with lower fluid intelligence. The pattern appears to be reversed in the context of activities presenting a low cognitive load. Similarly, Bonaiuto et al. (2016) validated their hypothesis that flow states are not affected by age or gender. However, these studies come from the broader flow literature. Additional research should be done to identify factors specifically affecting immersion–fluency, in addition to ruling out the effects of age.

7. Conclusion

7.1. Limitations and Future Research

7.1.1. Improving use of Subjective Measures of VR Psychological Experiences

A potential limitation of this study concerns the use of a single post-test questionnaire to evaluate pilot VR psychological experiences. Though this self-report format has a number of advantages, including demonstrating strong face validity in measuring the intended concepts, and serving as an efficient format for data collection and analysis, it has the disadvantage of requiring participants to answer retrospectively (van Baren & IJsselsteijn, 2004). A retrospective evaluation can introduce a recency bias, whereby participants' perceptions are more skewed towards events occurring at the end of the exercise. In the present study, pilot SA (i.e. information processing and comprehension) was evaluated as a knowledge state. In order to mitigate recency bias and allow for a more precise evaluation of the impacts of pilots' VR psychological experiences on their SA after each lag of flight, future iterations of this study should entail administration of the VR psychological experiences questionnaire alongside each SAGAT freeze. Moreover, adjustments might be made to the study to allow participants to complete questionnaire items by interacting with content in the simulated environment. For example, participants could be directed to provide verbal responses to text items appearing on screen; this approach was used by Kober and Neuper (2012). Integrating questionnaire items into VR content minimizes onus on individuals to make memory-based judgments, and it does not require interactions with external materials not relevant to the simulation, which might affect perceptions of interactivity and telepresence.

7.1.2. Integrating Objective Measures of Telepresence and Exploring Breaks in Presence

Future research in this area might also supplement subjective measures of VR psychological experiences with objective forms of assessment, which do not rely on participants' interpretation. Physiological measures, such as galvanic skin response (GSR) and electrocardiogram recordings (EKG), might be used to evaluate participants' physical responses to breaks in presence (BIPs) (Sanchez-Vives & Slater, 2005; Slater et al., 2006). BIPs refer to occurrences during exposure to a VR environment that cause an individual to become momentarily focused on stimuli from the external environment. Slater et al. (2006) used a combination of GSR and EKG to identify a physiological signature associated with the experience of a BIP. In addition, Kober and Neuper (2012) analyzed event-related brain potentials (ERP) from electroencephalographic (EEG) measures evoked by tones to assess telepresence experience in VR. The study used an oddball paradigm, whereby participants were exposed to a series of standard and deviant tones (administered 20% of the time at a much higher frequency), while moving through a simulated city. Participants were asked to evaluate their perceived telepresence at the end of the exercise and were then split into high- and low-telepresence groups. Different ERP components were compared between groups, and high-telepresence experience was found to be associated with lower amplitudes in frontal negative slow-waves. This area is associated with stimulus processing and attention allocation, suggesting that stronger perceived telepresence is related to increased commitment of attentional resources to VR content.

Future work should also consider the effects of system attributes on telepresence, as well as the incidence of BIPs. A higher graphics frame rate has been positively correlated with stronger experiences of telepresence (Sanchez-Vives & Slater, 2005). Similarly, lower levels of system latency in terms of the synchronization between head movements and the updating of displays have been associated with elevated levels of reported telepresence (Meehan et al., 2003). Additionally, Hou et al. (2016) compared the impact of two screen size conditions on participants' reported telepresence and other psychological factors during gameplay in a VR environment. The authors found that larger screen size had a positive influence on participants' sense of physical presence, as well as their mood and enjoyment.

When evaluating BIPs, the emotional valence of stimuli should also be considered. Since telepresence appears to emerge as predictions derived from a user's mental model are negotiated against actual outcomes in the virtual environment, content that compromises the availability or consistency of mental models may also result in BIPs (Liebold et al., 2017). Liebold et al. (2017) suggest that BIPs are likely to occur as the result of human orienting reflex/response, with more intense responses increasing the possibility of a BIP. In addition, emotional relevance of the stimuli can influence the intensity of orienting responses, with more negatively-associated, goal-pertinent stimuli provoking the strongest responses.

7.1.3. Improving the Research Design to Support Inferences of Causality

A limitation in the design of the present study is that only correlational relationships among factors can be inferred from the results of the analysis. Further work in this area could be enriched by exploring changes to the underlying methods, which would better enable investigation of causal relationships between pilots’ psychological experiences during VR exposure and their SA during flight. A robust causal study should demonstrate temporal precedence of causal factors, provide evidence for covariation of cause and effect, and rule out influences of extraneous variables. For example, a future study investigating the effects of BIPs on pilot SA during a VR flight might compare participants’ SA results during a simple flight exercise against their SA after they have been exposed to an interfering visual stimulus with either positive, negative, or neutral emotional valence. This stimulus should be an object or event that would be otherwise inconsistent with participants’ expectations for the task environment. Moreover, instead of emotional valence, treatment conditions could also compare effects of BIPs caused by negatively-associated stimuli engaging different sensory modalities on SA. For example, researchers might expose participants to a sudden loss of screen brightness or to an intrusive tone for a fixed duration of time. The order in which the control flight and treatment flight exercises are completed should be counterbalanced to minimize effects that exposure to the virtual environment might have on participants’ responses to the experimental stimulus. This arrangement would help to mitigate the influences of external variables when evaluating the effects of the study’s independent variables (i.e., interfering stimuli and emotional valence/sensory modality) on the dependent variables (i.e., SA and perceived presence), in addition to ensuring proper temporal ordering of the interfering stimulus and the SA task.

7.1.4. Increasing Sample Size

Another potential limitation of the present work was the small sample size of pilots tested, which may potentially compromise the reliability of results (i.e., the extent to which the current results can be reproduced under the same conditions). Effect sizes for the total effects of relationships were small and medium in value according to guidelines by Cohen (1988). Moreover, a potential sample size issue was noted in the fact that the direct relationship between experience and SA Level I slipped below the threshold for significance, $p = .06$, when evaluated in the larger combined model of all VR psychological experience and SA factors. It is therefore recommended that these findings be validated with a larger data set.

7.2. Summary and Implications

The main objective of this research was to validate and quantify the effects of telepresence, interactivity, and immersion on pilot SA at the first and second level. Findings from the present study suggest that telepresence supports SA in VR environments at both levels. Though interactivity only directly influenced SA at the first level, it also had a significant supportive effect. Exploring relationships of telepresence and interactivity to working memory may help to account for the positive effects of these psychological experiences on pilot SA during VR flight tasks. High levels of interactivity suggest that the user is focused on elements of the virtual environment and effectively filtering out external noise stimuli. In contrast, telepresence is thought to develop through the formation of more complex cognitive representations, which may provide a foundation for achievement of higher-level SA. Despite the wealth of research suggesting that flow states have a positive effect on performance during cognitively demanding tasks, the effects of immersion–fluency on SA were shown to be non-significant at both levels. These results may potentially be explained by attributes of the virtual environment being overly fascinating, complex, or unintuitive, such that they draw the user's attention away from the primary task (Bian et al., 2018).

As a secondary line of inquiry, the effects of age on the aforementioned three VR psychological experience constructs were also explored to shed greater insight on a range of accounts from the literature citing positive, negative, and non-significant relationships. Results from the path analysis revealed an incidence of Simpson's paradox when evaluating a potential direct relationship between age and telepresence, suggesting the influence of an unidentified third variable. Additional research is needed to better understand the factors influencing telepresence. Age was shown to have a significant positive effect on participants' perceived interactivity. A potential explanation for this finding may be related to the flight skill that older pilots brought to the exercise. For example, studies by Ghani and Deshpande (1994) and Novak et al. (2000) demonstrate direct relationships between skill and perceived control. Finally, as was expected, no significant relationship was observed between age and immersion–fluency.

The results of this research provide meaningful contributions that can be used towards improving the design of VR flight simulators for the purpose of training and assessment of pilot SA. The present study demonstrates that pilot's psychological experiences of telepresence and interactivity during exposure to VR flight simulators have serious implications in terms of their ability to achieve SA at both the level of information processing and comprehension. Moreover, telepresence experience is highly dependent on a pilot's perception of system interactivity. Development efforts around VR-based cognitive training tools for GA pilots should explore strategies for mitigating BIPs, or occurrences that cause an individual to become focused on stimuli from the external environment. Such strategies should explore effects of both the cognitive demand of the tasks being performed, as well as the integration of sensory information communicated by the system, on the incidence of BIPs. Avoiding events that compromise the continuity and cohesiveness of a pilot's experience of telepresence are essential for ensuring the integrity of the VR training environment and enhancing the likelihood of knowledge transfer between simulated and real-world contexts.

8. References

Baddeley, A. D. (2000). The episodic buffer: A new component of working memory?. *Trends in Cognitive Sciences*, *4*(11), 417-423.

Baddeley, A. D., and Hitch, G. (1974). Working memory. In G.H. Bower (Ed.), *The psychology of learning and motivation: Advances in research and theory* (Vol. 8, pp. 47–89). New York: Academic Press.

Bangay, S., and Preston, L. (1998). An investigation into factors influencing immersion in interactive virtual reality environments. *Stud. Health Technol. Inform.*, *58*, 43-51.

Bian, Y., Yang, C., Zhou, C., Liu, J., Gai, W., Meng, X., Tian, F., and Shen, C. (2018). Exploring the weak association between flow experience and performance in virtual environments. In *Proc. of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, 21-26 April 2018*: ACM.

Biocca, F. (1997). The cyborg's dilemma: progressive embodiment in virtual environments. *Journal of Computer-Mediated Communications, 3*(2).

Bolstad, C.A., Endsley, M.R., Costello, A.M., and Howell, C.D. (2010). Evaluation of Computer-Based Situation Awareness Training for General Aviation Pilots. *The International Journal of Aviation Psychology, 20*(3), 269-294.

Bolstad, C.A., and Hess, T.M. (2000). Situation Awareness and Aging. In M.R. Endsley and D.J. Garland (Eds.), *Situation Awareness Analysis and Measurement* (pp. 277-302). Boca Raton, FL: CRC Press.

Bonaiuto, M., Mao, Y., Roberts, S., Psalti, A., Ariccio, S., Ganucci Cancellieri, U., and Csikszentmihalyi, M. (2016). Optimal Experience and Personal Growth: Flow and the Consolidation of Place Identity. *Frontiers in Psychology*, 7, 1654. doi: 10.3389/fpsyg.2016.01654.

Buttussi, F., and Chittaro, L. (2018). Effects of Different Types of Virtual Reality Display on Presence and Learning in a Safety Training Scenario. *IEEE Transactions on Visualization and Computer Graphics, 24*(2), 1063-1076.

Bystrom, K., Barfield, W., and Hendrix, C. (1999). A conceptual model of the sense of presence in virtual environments. *Presence, 8*(2), 241–244.

Christopher, K. (2017). Report on the COPA 2017 Membership Survey. Ottawa, ON: Canadian Owners and Pilots Association.

Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences*. New York, NY: Routledge Academic.

Corriveau Lecavalier, N., Ouellet, E., Boller, B., and Belleville, S. (2018). Use of immersive virtual reality to assess episodic memory: A validation study in older adults. *Neurological Rehabilitation*, DOI: 10.1080/09602011.2018.1477684

Csikszentmihalyi, M. (1990). *Flow: The Psychology of Optimal Experience*. New York, New York: Harper & Row Publishers Inc.

Damasio, A.R. (1998). Investigating the biology of consciousness. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, 353, 1879-1882.

Dörr, K., Schiefele, J., and Kubbat, I. W. (2000). Virtual Cockpit Simulation for Pilot Training. *In Proc. Of the Annual North Atlantic Treaty Organization (NATO) Research and Technology Organization Human Factors and Medicine Panel (RTO HFM) Workshop on "What is essential for virtual reality systems to meet* *military human performance goals?", The Hague, Netherlands, 13-15 April 2000*: RTO MP-058.

Endsley, M.R. (1990). Predictive utility of an objective measure of situation awareness. *In Proc. of the Human Factors Society 34th Annual Meeting, Santa Monica, USA, October 1990.*

Endsley, M.R. (1995a). Toward a theory of situation awareness in dynamic systems. *Human Factors Journal*, 37(1), 32-64.

Endsley, M.R. (1995b). A taxonomy of situation awareness errors, human factors in aviation operations. In R. Fuller, N. Johnston, and N. McDonald (Eds.), *Human Factors in Aviation Operations* (pp. 287-292). Aldershot, England: Avebury Aviation, Ashgate Publishing Ltd.

Endsley, M.R. (1995c). Measurement of Situation Awareness in Dynamic Systems. *Human Factors, 37*(1), 65-84.

Endsley, M.R., and Jones, D. G. (2012). *Designing for Situation Awareness: An Approach to User-Centred Design (Second Edition)*. Boca Raton, FL: CRC Press, Taylor and Francis Group.

Engeser, S., and Rheinberg, F. (2008). Flow, performance, and moderators of challenge-skill balance. *Motivation and Emotions, 32*(3), 158-172.

Engeser, S., Rheinberg, F., Vollmeyer, R., and Bischoff, J. (2005). Motivation, Flow-Erleben und Lernleistung in universitären Lernsettings. [Motivation, flow experience, and performance in learning settings at university]. *Zeitschrift für Pädagogische Psychologie*, 19, 159–172.

Federal Aviation Administration. (2018). 2018 *Active Civil Airmen Statistics*. Washington, DC: U.S. Department of Transportation.

Ghani, J., and Deshpande, S.P. (1994). Task characteristics and the experience of optimal flow in human-computer interaction. *The Journal of Psychology Interdisciplinary and Applied, 128*(4), 381-391.

Goldberg, I. I., Harel, M., and Malach, R. (2006). When the brain loses its self: Prefrontal inactivation during sensorimotor processing. *Neuron, 50*, 329–339.

Gorman, P.J., Meier, A.H., and Krummel, T.M. (1999). Simulation and Virtual Reality in Surgical Education: Real or Unreal. *The Archives of Surgery, 134*(11), 1203–1208.

Hair, J.F., Sarstedt, M., Hopkins, L., and Kuppelwieser, V.G. (2014). Partial least squares equation modeling (PLS-SEM): An emerging tool is business research. *European Business Review, 26*(2), 106-121.

He, Z., Zhu, F., Perlin, K., and Ma, X. (2018). Manifest the Invisible: Design for Situational Awareness of Physical Environments in Virtual Reality. Submitted for publication **arXiv:1809.05837v1**.

Heeter, C. (2002). Being There: The Subjective Experience of Presence. *Presence: Virtual and Augmented Reality, 1*(2), 262-271.

Hess, T.M. (1990). Aging and schematic influences on memory. In T.M. Hess (Ed.), *Aging and cognition: Knowledge organization and utilization* (pp.93-160). Amsterdam: North-Holland.

Hoffman, D. L., and Novak, T. P. (1996). Marketing in hypermedia computer-mediated environments: Conceptual foundations. *Journal of Marketing, 60*, 50–68.

Hou, J., Nam, Y., Peng, W., and Lee, K.M. (2012). Effects of screen size, viewing angle, and players' immersion tendencies on game experience. *Computers in Human Behavior, 28*(2), 617-623.

Hsu, C. (2010). Exploring the player flow experience in e-game playing. *International Journal of Technology and Human Interaction, 6*(2), 47+.

International Society for Presence Research. (2000). The Concept of Presence: Explication Statement. Retrieved July 24, 2019 from https://ispr.info/

Jung, D., Jo, S., and Myung, R. (2008). A Study in Relationships Between Situation Awareness and Presence that Affect Performance on a Handheld Game Console. *In Proc. of the 2008 International Conference on Advances in Computer Entertainment Technology, Yokohama, Japan, 3-5 December 2008:* ACE-08.

Kim, T., and Biocca, F. (1997). Telepresence via television: Two dimensions of telepresence may have different connections to memory and persuasion. *Journal of Computer-Mediated Communication*, 3(2).

Kiousis, S. (2002). Interactivity: A Concept Explication. *New Media and Society, 4*(3), 355-383.

Kober, S.E., and Neuper, C. (2012). Using auditory event-related EEG potentials to assess presence in virtual reality. *International Journal of Human-Computer Studies, 70*(9), 577-587.

Kock, N. (2018). WarpPLS User Manual: Version 6.0. Lardeo, TX: ScriptWarp Systems.

Laptaned U. (2006). Situation Awareness in Virtual Environments: A Theoretical Model and Investigation with Different Interface Designs. *In Proc. of the 9th IASTED*

International Conference Computer and Advanced Technology in Education, Peru, 2006.

Liebold, B., Brill, M., Pietschmann, D., Schwab, F., and Ohler, P. (2017). Continuous Measurement of Breaks in Presence: Psychophysiology and Orienting Responses. *Media Psychology, 20*(3), 477-501.

McGreevy, M. W. (1992). The presence of field geologists in Mars-like terrain. *Presence: Teleoperators and Virtual Environments, 1*(4), 375–403.

McMillan, S.J., and Hwang, J. (2002). Measures of perceived interactivity: an exploration of the role of direction of communication, user control, and time in shaping perceptions of interactivity. *Journal of Advertising*, *31*(3).

Meehan, M., Razzaque, S., Whitton, M.C., and Brooks, F.P. (2003). Effect of latency on presence in stressful virtual environments. *In Proc. of the IEEE Virtual Reality Annual International Symposium, Los Angeles, California, 22-26 March 2003.*

Mestre, D., and Vercher, J. (2011). Immersion and presence. In P. Fuchs, G. Moreau, and P. Guitton (Eds.), *Virtual Reality: Concepts and Technologies* (pp. 93-100). Presses des Mines, Paris, France: CRC Press.

Michailidis, L., Balaguer-Ballester, E., and He, X. (2018). Flow and Immersion in Video Games: The Aftermath of a Conceptual Challenge. *Front. Psychol.* https://doi.org/10.3389/fpsyg.2018.01682

Minksy, M. (1980, June). Telepresence. OMNI Magazine, p.45-52. Retrieved from https://web.media.mit.edu/~minsky/papers/Telepresence.html.

Mütterlein, J. (2018). The Three Pillars of Virtual Reality? Investigating the Roles of Immersion, Presence, and Interactivity. *In Proc. of the 51st Annual Hawaii*

International Conference on System Science, Honolulu, Hawaii, 3-6 January 2018: HICSS-51.

Mütterlein, J., and Hess, T. (2017). Immersion, Presence, Interactivity: Towards a Joint Understanding of Factors Influencing Virtual Reality Acceptance and Use. *In. Proc. of the 23rd Americas Conference on Information Systems, Boston, MA, 10-12 August 2017:* AMCIS-23.

Ninaus, M., Pereira, G., Stefitz, R., Prada, R., Paiva, A., Neuper, C., and Wood, G. (2015). Game elements improve performance in a working memory training task. *International Journal of Serious Games, (2)*1, 3-16.

Norman, D. A. (1988). The Design of Everyday Things. New York: Doubleday.

Novak, T. P., Hoffman, D. L., and Yung, Y.F. (2000). Measuring the customer experience in online environments: A structural modeling approach. *Marketing Science, 19*(1), 22–42.

Payne, B. R., Jackson, J. J., Noh, S. R., and Stine-Morrow, E. A. (2011). In the zone: flow state and cognition in older adults. *Psychology and Aging, 26*(3), 738–743.

Pearl, J. (1999). Simpson's Paradox: An Anatomy. UCLA Cognitive Systems Laboratory, Technical Report.

Prince, C., and Salas, E. (1997). Situation assessment for routine flight and decision making. *International Journal of Cognitive Ergonomics, 1*, 315–324.

Prothero, J. D., Hoffman, H. G., Parker, D. E., Furness, T. A., and Wells, M. J. (1995). Foreground/Background Manipulations Affect Presence. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 39*(21), 1410–1414. https://doi.org/10.1177/154193129503902111

Psotka, J. (1995). Immersive training systems: Virtual reality and education and training. *Instructional Science*, *23*(5-6), 405-431.

Rawlinson, T.G., Lu, S., and Coleman, P. (2012). Individual Differences in Working Memory Capacity and Presence in Virtual Environments. *In Proc. of Advances in Brain Inspired Cognitive Systems: 5th International Conference, Shenyang, China, 11-14, July 2012:* BICS-2012.

Regenbrecht, H., Schubert, T., and Friedmann, F. (1998). Measuring the Sense of Presence and its Relations to Fear of Heights in Virtual Environments. *International Journal of Human-Computer Interaction, 10*(3), 233-249.

Rheinberg, F., Vollmeyer, R., and Engeser, S. (2003). Die Erfassung des Flow-Erlebens [The Assessment of Flow Experience]. In J. Stiensmeier-Pelster, & F. Rheinberg (Eds.), Diagnostik von Selbstkonzept, Lernmotivation und Selbstregulation [Diagnosis of Motivation and Self-Concept]. Göttingen: Hogrefe.

Riley, J.M., Kaber, D.B., and Draper, J.V. (2004). Situation Awareness and Attention Allocation Measures for Quantifying Telepresence Experiences in Teleoperation. *Human Factors and Ergonomics Manufacturing, 14*(1), 51-67.

Salthouse, T.A. (1984). Effects of age and skill in typing. *Journal of Experimental Psychology: General*, *113*, 345-371.

Salthouse, T.A. (1991). Working memory as a processing resource in cognitive aging. *Developmental Review, 10*(1), 101-124.

Sanchez-Vives, M.V., and Slater, M. (2005). From Presence to Consciousness Through Virtual Reality. *Nature Reviews Neuroscience*, *6*, 332-339.

Schubert, T., Friedmann, F., and Regenbrecht, H. (1998). Embodied Presence in Virtual Environments. In R. Payton and I. Nielson (Eds.), *Visual Representations and Interpretations* (pp. 269-278). London, UK: Springer.

Schubert, T., Friedmann, F., and Regenbrecht, H. (2001). The Experience of Presence: Factor Analytic Insights. *Presence, 10*(3), 266-281.

Schüler, J. (2007). Arousal of Flow Experience in a Learning Setting and Its Effects on Exam Performance and Affect. *Zeitschrift für Pädagogische Psychologie*, 21(3), 217-227.

Sheridan, T.B. (1992). Musings on telepresence and virtual presence. *Presence: Teleoperators and Virtual Environments*, *1*(1), 120-125.

Shernoff, D., and Csikzentmihalyi, M. (2009). Flow in schools revisited: Cultivating engaged learners and optimal learning environments. In M. Furlong., R. Gilman, and S. Heubner (Eds.), *Handbook of Positive Psychology in Schools (2nd Edition)* (pp. 211-226). New York, NY: Taylor and Francis Group.

Siebert, J., and Shafer, D.M. (2017). Control mapping in virtual reality: effects on spatial presence and controller naturalness. *Virtual Reality, 22*(1), 79-88.

Slater, M., and Wilbur, S. (1997). A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments. *Presence, Teleoperators, and Virtual Environments, 6*(6), 603-616.

Slater, M., Guger, C., Edlinger, G., and Leeb, R. (2006). Analysis of physiological responses to a social situation in an immersive virtual environment. *Presence: Virtual and Augmented Reality, 15*(5), 553-569.

Steur, J. (1992). Defining Virtual Reality: Dimensions Determining Presence. *Journal of Communication, 42*(4), 73-93.

Transportation and Safety Board of Canada. (2018). Statistical Summary: Air Occurrences in 2017. Retrieved from http://www.bst-tsb.gc.ca/eng/stats/aviation/2017/ssea-ssao-2017.asp

van Baren, J., and IJsselsteijn, W. (2004). Measuring Presence: A Guide to Current Measurement Approaches (OmniPres project IST-2001-39237). Eindhoven, Netherlands: Eindhoven University of Technology. Retrieved from International Society for Presence Research website http://www8.informatik.umu.se/~jwworth/PresenceMeasurement.pdf

Van Benthem, K., Herdman, C. M., Brown, M., & Barr, A. (2011). The Relationship of age, experience and cognitive health to private pilot situation awareness performance. Proceedings of the 16th International Symposium on Aviation Psychology, 2–5.

von der Pütten. A.M., Klatt, J., Broeke, S.T., McCall, R., Krämer, N.C., Wetzel, R., Blum, L., Oppermann, L., and Klatt, J. (2012). Subjective and behavioural presence measurement and interactivity in the collaborative augmented reality game TimeWarp. *Interacting with Computers*, 24, 317-325.

Weibel, D., Wissmath, B., Habegger, S., Steiner, Y., and Groner, R. (2008). Playing online games against computer- vs. human-controlled opponents: Effects on presence, flow, and enjoyment. Computers in Human Behavior, 24, 2274-2291.

Wetzels, M., Odekerken-Schroder, G., and van Oppen, C. (2009). Using PLS path modeling for assessing hierarchical construct models: Guidelines and empirical illustration. *MIS Quarterly, 33*(1), 177-196.

Wirth, W., Hartman, T., Böcking, S., Vorderer, P., Klimmt, C., Schramm, H., Saari, T., Laarni, J., Ravaja, N., Gouveia, F.R., Biocca, F., Sacau, A., Jäncke, L., Baumgartner, T., and Jäncke, P. (2003). Constructing Presence: Towards a Two-Level model of the Formation of Spatial Presence Experience. Report to the European Community, Project Presence: MEC (IST-2001-37661).

Witmer, B. G., and Singer, M. J. (1998). Measuring Presence in Virtual Environments: A Presence Questionnaire. *Presence: Teleoperators and Virtual Environments,* 7(3), 225-240.

Appendices

Appendix A. Certification of Institutional Ethics Clearance

Office of Research Ethics
503 Robertson Hall | 1125 Colonel By Drive
Ottawa, Ontario K1S 5B6
613-520-2600 Ext: 4085
ethics@carleton.ca

CERTIFICATION OF INSTITUTIONAL ETHICS CLEARANCE

The Carleton University Research Ethics Board-B (CUREB-B) at Carleton University has renewed ethics clearance for the research project detailed below. CUREB-B is constituted and operates in compliance with the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (TCPS2).

Title: The Usefulness of Cognitive Health Screening for Predicting Risk in General Aviation [Kathleen Van Benthem]

Protocol #: 108569

Principal Investigator: Dr. Kathleen Van Benthem

Department and Institution: Faculty of Arts and Social Sciences\Cognitive Science (Institute of), Carleton University

Project Team (and Roles): Dr. Kathleen Van Benthem (Primary Investigator)
Caidence Paleske (Student Researcher)
Cassandra Ommerli (Other)
Jinous Mirzaagha (Other)
Sarah Enouy (Other)
Adam Fraser (Other)
Caitlin Shanahan (Other)
Dr. Chris Herdman (Research Supervisor)

Funding Source (If applicable):

Effective: **February 21, 2019** Expires: **February 29, 2020**.

Please ensure the study clearance number is prominently placed in all recruitment and consent materials: CUREB-B Clearance # 108569.

<u>Restrictions:</u>

This certification is subject to the following conditions:

1. Clearance is granted only for the research and purposes described in the application.

2. Any modification to the approved research must be submitted to CUREB-B. All changes must be approved prior to the continuance of the research.

3. An Annual Application for the renewal of ethics clearance must be submitted and cleared by the above date. Failure to submit the Annual Status Report will result in the closure of the file. If funding is associated, funds will be frozen.

4. A closure request must be sent to CUREB-B when the research is complete or terminated.

5.During the course of the study, if you encounter an adverse event, material incidental finding, protocol deviation or other unanticipated problem, you must complete and submit a Report of Adverse Events and Unanticipated Problems Form, found here: https://carleton.ca/researchethics/forms-and-templates/

6. It is the responsibility of the student to notify their supervisor of any adverse events, changes to their application, or requests to renew/close the protocol.

7. Failure to conduct the research in accordance with the principles of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans 2nd edition* and the *Carleton University Policies and Procedures for the Ethical Conduct of Research* may result in the suspension or termination of the research project.

Upon reasonable request, it is the policy of CUREB, for cleared protocols, to release the name of the PI, the title of the project, and the date of clearance and any renewal(s).

Please email the Research Compliance Coordinators at ethics@carleton.ca if you have any questions.

CLEARED BY: **Date: February 21, 2019**

[redacted]

Bernadette Campbell, PhD, Chair, CUREB-B

[redacted]

Natasha Artemeva, PhD, Vice-Chair, CUREB-B

Appendix B. Participant Consent Form

Study: The Usefulness of Cognitive Health Screening for Predicting Risk in General Aviation
Faculty Sponsor: Dr. Chris M. Herdman, Department of Psychology, Carleton University, tel. 613-520-2600 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 can determine whether you wish to participate in the study. This research was cleared by the Carleton University Research Ethics Board – B February 2018. Please ask the researcher to clarify any concerns that you may have after reading this form.

Project Protocol Clearance #: *108569*
Expiry: *December 31, 2019.*

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	Email	Phone
Kathy Van Benthem	Postdoctoral Fellow	kathy.vanbenthem@carleton.ca	520-2600 x.2487
Caidence Paleske	Masters Student	caidencepaleske@cmail.carleton.ca	520-2600 x.2487
Connor Branch	Honours Student	connorbranch@cmail.carleton.ca	520-2600 x.2487
Alicia Fu	Honours Student	AliciaFu@cmail.carleton.ca	520-2600 x.2487
Alicia Krolak	Independent Study	aliciaKrolak@cmail.carleton.ca	520-2600 x.2487

Other Contacts: Should you have any ethical concerns regarding this study, please contact:

Name	Contact Info.
Carleton University Research Office	ethics@carleton.ca
Dr. Andy Adler, Chair, Carleton University Research Ethics Board - B	ethics@carleton.ca

If you have any ethical concerns with the study, please contact Dr. Andy Adler, Chair, Carleton University Research Ethics Board-B and the Carleton University Research Compliance Office (by phone at 613-520-2600 ext. 4085 or via email at ethics@carleton.ca)

Purpose: The purpose of this study is to determine the usefulness of cognitive health screening in predicting risk in general aviation. Measures collected during a simulated search and rescue (SAR) training exercise will be compared with performance on two cognitive screening tests (see page 4:A-C for an illustration of these tests). The general aviation performance measures include critical incidents, flight path deviations, memory for tasks, and situation awareness. We also track pilot mental workload using self-report, and biometrics, such as heart rate--using a wristband, and electroencephalographic (EEG) response--using a non-invasive EEG device (see page 5:D & E for a picture of these devices). Part of the time you will fly the SAR training exercise with an electronic navigational aid (see page 5:F), as we are interested in the impact of electronic navigational flight aids in general aviation.

Participants Eligibility
You are eligible to participate if you are a licensed aeroplane pilot with a valid medical certification, and you are at least 18 years of age. You should also have flown as pilot-in-command in the past 24-months. If you have a student permit you may be eligible to participate if you have "solo" and cross-country flight experience.

Study Tasks:

Session 1 Tasks: After signing the consent form we will ask you to complete a demographic questionnaire e.g., age, experience as a pilot. If you are not familiar with the ForeFlight navigational aid, you will then be asked to watch a brief instructional video like the online training video we suggested you watch before coming in for session one. You will have the opportunity to become familiar flying the Cessna 172 simulator and will fly basic maneuvers around the Ottawa airspace (while using the electronic navigational flight aid). We will review the use of the ForeFlight electronic navigational aid with you before you fly a custom-designed SAR training exercise in local airspace.

In the SAR training exercise, you will:

- fly circuits and a cross-country route that includes taking-off,
- lying at a pre-determined altitude and executing turns and landings.
- There will be other "virtual" aircraft flying in the airspace that communicate air-to-air.
- You will be briefed on the search and rescue training exercise.
- e will ask you to wear a wireless EEG headset and a biometric wristband during the 60-minute search and rescue flight (see page 5).

Situation Awareness: At two points during the SAR training exercise we will pause the scenario and ask you questions pertaining to situation awareness. These questions are related to the location and details of other aircraft and the location and flight details of your own aircraft. You will also be asked to rate yourself on a situation awareness self-rating scale.

Prospective Memory Task: During the search and rescue training exercise you will monitor a small screen in the cockpit where arrows will be periodically displayed. You will be asked to make a short radio call when you observe a right-facing arrow.

Mental Workload Measures: Mental workload will be measured during the search and rescue training exercise. The first method uses a peripheral detection task (PDT), where you will depress a small switch on the Cessna yoke whenever you hear a "tone" played via a speaker. Workload will also be measured via two biometric devices: the Empatica E4 wristband and the EPOC+ electroencephalography (EEG) headset by EMOTIV (see page 5). The Empatica E4 wristband is a wearable wireless device designed for continuous, real-time data collection. It has a light sensor that measures blood volume pulse from which heart rate and heart rate variability may be derived. It also measures electrodermal activity (sweat) via a galvanic skin response sensor. This band also uses a sensor to read your skin temperature. The EPOC+ EEG headset detects electrical potentials at the scalp from 14 felt sensors, it is harmless and non-invasive (see page 5). Finally, you will be asked to report on your perception of the mental demands of the flight during the two pauses in the scenario.

Session 2 Tasks: On your second visit to the ACE Lab we will ask you to complete a questionnaire about your experience with electronic navigational flight aids. You will then complete two 20-minute cognitive screening tests (see page 4).

- The first test is the CanFly (our custom virtual reality test for pilots. The CanFly requires you to fly a short flight plan where you maintain situation awareness and respond to questions about other aircraft in the airspace. The CanFly collects data on basic aviator skills such as maintaining a prescribed altitude and remembering details of the flight environment. The CanFly test uses the Oculus Rift to deliver the graphics for the flight. The flight controls are managed through a simple custom Cessna 172 simulator. Before and after the CanFly virtual reality flight test, you will answer a few questions pertaining to how you are feeling.

- The second test is the MATB-II (a 20-minute computerized designed by NASA). The MATB-II test requires you to monitor gauges, track a shape, and change radio frequencies in response to messages you will hear via a headset. You will be asked to complete three five-minute sessions of the MATB-II that represent easy, medium, and difficult conditions (but not necessarily presented in that order)

Locale, Duration, and Affiliation: Testing will be completed in two phases: both will take place in VSIM 1214 at Carleton University. The total time to complete this study is four-hours and is spread over two sessions. This research is not affiliated with any device, flying club, flight school or aviation program.

Compensation: You will not be compensated monetarily. We compensate you for your time by paying for your parking on the Carleton Campus and by providing you with refreshments each session you attend.

Potential Risks/Discomfort: There are no potential psychological risks associated with participation in this experiment. Please note that your performance in the full-scale Cessna simulator does not provide an indication of your skills as a pilot. However, if you feel anxious and/or uncomfortable about your performance in this experiment or feel effects of the Cessna 172 or VR simulation, please bring your concerns to the researcher's attention and we will remove the device right away. After five minutes you can decide if you would like to continue the study. Any risk to well-being caused by the virtual reality device is considered minimal and usually disappears after a minute or two. Previous research using VR and a Cessna simulator in our lab with 40 participants did not result in any participants feeling ill after the simulated flight.

Anonymity/Confidentiality: All data collected in this experiment will be kept anonymized through the assignment of a coded number to the data. This data is securely stored on a local computer for a maximum of ten years. Similarly, 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 and may therefore be presented and/or published. You can not be identified in any reports produced from this study. Disclaimer: *Your data will be stored and protected by researchers at the ACE Lab in a secure server in our locked facility at Carleton University but may be disclosed via a court order or data breach.*

Right to Withdraw/Omit: You have the right to withdraw at any time during session one or two without penalty. Before you leave on session two, you have the right and ability to withdraw from this experiment without penalty (e.g., refreshments are still offered to you and you do not have to repay the value of the parking pass), and all your data will be deleted. However, because the data collected is anonymized, after you leave session two you will not be able to withdraw from the study. Your participation in this experiment is completely voluntary.

____

I have read the above description of the study examining the usefulness of a cognitive assessment battery in predicting general aviation flight performance.

Name: ____________________________
Date: _____________________________
Signature: _________________________
Witness: __________________________

3