Exploratory study of virtual reality flight training device for upset prevention and recovery training

Filip Florek

Cranfield University, UK

SUMMARY

This exploratory study aimed to establish whether a virtual reality (VR) enabled flight training device will provide effective Upset Prevention and Recovery Training (UPRT) to ensure equivalent safety with the Flight Simulation Training Device (FSTD) considering pilot situational awareness competency. It was achieved by determining the effects of the pilot's presence, task-related stress and cybersickness on situational awareness during upset prevention, and, if necessary, recovery and by assessing pilot acceptance of VR-enabled flight training device in UPRT. No evidence has been found that situational awareness was negatively affected by exposure to VR, with certain reactions to stimuli degradation, i.e., flight upset in Instrument Meteorological Conditions (IMC) resulting in a predictable outcome of increased attentional demand.

KEYWORDS

Serious Games for Training, Simulation, Situation awareness, Virtual Reality, Training.

Introduction

Many sources (Airbus, 2022; Boeing, 2021) cite Loss of Control In-flight (LOC-I) as a leading causal factor in fatal accidents in civil aviation over the last 30 years. In recent decades technological advancements like flight envelope protection and advanced upset warning systems allowed to reduce LOC-I fatal accident rates by nearly 90% (Airbus, 2022), but still high-profile accidents like Air France (Flight 447) and Colgan Air (Flight 3407) undermined the mitigating factors and proved that that flight automation alone will not resolve the issue (Richards et al., 2012). In both cases, the accident report articulates the complexity of contributing factors leading to airplane aerodynamic stall and pilot's cognitive performance degradation, especially in situational awareness and as result inadequate response by pilots (BEA, 2012; NTSB, 2009).

Aircraft upset describes an in-flight state in which an aircraft exceeds structural parameters of the airframe (ICAO, 2014b). These disturbances may result in a stall, spin, or over-limit angle of attack (Brooks & Ransbury, 2019). In an unexpected airplane upset event, interrelated factors in aircraft handling, inability to comprehend unfamiliar stimuli, and the psychological stressors of surprise, startle, and fear can combine to create compound threats (Brooks & Ransbury, 2019). Upset recognition and recovery skills require timely and rapid application of corrective inputs, skills that needs to be trained. Upset prevention and recovery training (UPRT) become a focus area of airline operations and training (Rogers et al., 2009). The use of high-end FSTDs for the delivery of UPRT during flight training, complements the application of knowledge and techniques introduced through on-aeroplane UPRT (ICAO, 2014b). The operational potential of FSTD allows for training in upset areas, i.e., low, or very high altitude or in adverse weather conditions, that can be deemed unsafe or impracticable in real aircraft (ICAO, 2015; Miglior, 2014). Growing evidence (Leland et al., 2009; Rogers et al., 2009) suggests that low-cost simulation systems deliver comparable training transfer in UPRT at the fraction of the cost of the FFS. It is essential to establish whether the design of a VR-enabled flight training device will provide effective UPRT training to ensure equivalent safety and effectiveness with the FSTD considering pilot core competencies. This exploratory research addresses two aims. First to determine the effects of the pilot's presence, task-related stress and cybersickness on situational awareness during upset prevention, and, if necessary, recovery. Second to assess pilot acceptance of VR-enabled flight training device in UPRT. The quality of upset recovery training, including the delivery method, the training content, and the training transfer are excluded from this study.

To perform the tasks of upset prevention and recovery, a flight crew needs to deploy several competencies (ICAO, 2014b). Situational awareness and decision-making are critical competencies during the prevention phase, while the application of procedures and aeroplane flight path management - manual control (ICAO, 2014b) are the most critical competencies during recovery from an upset condition. Therefore, the key criterium in the selection of the evaluative scenarios would be an application of situational awareness in prevention and manual control for recovery phases. This competency is well suited to an evaluation in the VR-based environment considering technological limiting factors. Loss of situational awareness among pilots is a well-researched topic (Endsley, 1995; Endsley et al., 2000; Endsley & Jones, 2016; Jones & Endsley, 1996; Stark et al., 2001; van de Merwe et al., 2012). Pilot's errors on the flight deck are typically attributed to disruptions in the decision-making process, however, according to Endsley (1995), it is not the response to the situation but limited or impaired perception and comprehension of the situation - the actual causal factors of accidents. Decision-making relies on situational awareness as a critical factor (Ommerli, 2019), achieving situational awareness is cognitively demanding and it is central to task performance (Endsley, 1995).

One essential part of the UPRT is the skill of recognising all required stimuli and processing the information in upset conditions, and not only from memorised procedures (Brooks & Ransbury, 2019). Pilot situational awareness can be impaired when exposed to physical (i.e., vibration, temperature, lighting, fatigue) or psychological stress (i.e., workload, time pressure, fear, or uncertainty) (Hockey, 1986). There are few documented symptoms of stress factors influencing situational awareness, like a narrow field of attention (focus on a limited number of central cues), cognitive tunnel vision (sampling only obvious or probable sources of information) and premature closure (deciding without exploring all options) (Endsley, 1995). These symptoms affect predominantly the early stage of the decision-making process involving perception, as a result, the assessment of the situation and projection of near-future events (i.e., how the situation may evolve) are impacted by limited recognition of the elements and attributes of a system. Application of situational awareness in prevention and manual control for recovery phases is well suited to evaluate in the VR-based environment considering technological limiting factors.

The impression of being in the virtual environment is a state of psychological awareness commonly referred to as presence (Slater & Wilbur, 1997). In other words, the greater degree of presence, the more likely that humans exposed to VR will perceive the environment and react in a manner like their behaviour in the real world (Slater & Wilbur, 1997). According to Steuer and Reeves (1992) presence is one of the key defining features of virtual environments. The relationship between situational awareness and presence is well documented, and studies (He et al., 2018; Jung et al., n.d.; Prothero et al., 2016) report a significant positive association between these two constructs. Physical and visual motion in a simulator can cause a side-effect known as motion sickness, commonly referred to in this context as simulator sickness, with symptoms including visual disturbances, a decline in hand-eye coordination and gastrointestinal manifestations (Webb, 2010). Additional studies indicate also high severity of those symptoms with delayed effects (Kolasinski &

Gilson, 1998). In the context of virtual reality, simulator sickness is referred to as cyber sickness. Saredakis et al (2020) report that the key impactors on cybersickness are visual stimulation i.e., the content presented in VR, resolution and refresh rate, exposure time and level of locomotion.

Method

Within-subjects nonexperimental design was used. All recruited participants (n = 11) selected for the study were professional pilots with a background in military aviation with a total flying experience mean value of 1377 flight hours (SD = 1286). No power calculations for the sample were conducted due to the exploratory nature of the study.

The simulation software used in this study is based on a commercial off-the-shelf (COTS) X-Plane 11 system. Robin DR401 CDI with Garmin G1000 Electronic Flight Instrument System a general aviation type aeroplane was used during simulated upset scenarios The aircraft selection was a consequence of matching the aircraft as close as possible to the models used in the pilot's training represented by participants' sample. A self-assembled, fixed-base, VR FTD hardware platform was used in the study and consisted of the flight stick and throttle quadrant, Rudder pedals, and VR head-mounted display. HP Reverb G2 VR head-mounted display set has built-in audio capabilities to ensure that the auditory cues were provided during the flight scenarios.

UPRT scenarios emphasize the need for the pilot to maintain situational awareness to recognize a divergence from nominal conditions as early as possible and immediately take corrective action including managing the energy, arresting the flight path divergence, and recovering to a stabilized flight path (ICAO, 2014a). The ICAO (2014a) recommends 16 training that grouped by upset-inducing topics, with each topic consisting of the exercise conditions, training description and rationale. Four selected scenarios for the study (see **Error! Reference source not found.**) focused on maintaining situational awareness to immediately take corrective action in case of recognised upset by manipulation of the control surfaces and the throttle to maintain the aircraft attitude and correct, if necessary, to return the aircraft to a stabilized flight path.

Table 1: Four selected ICAO flight upset scenarios

Scenario/ Rationale	Task
S1. High-altitude upset with environmental factors as a causal factor	Task: Change altitude while maintaining airspeed.
S2. Clean configuration approach-to-stall (high altitude)	Task: Maintain altitude. Reduce thrust to less than adequate. Recognize the stall warning and perform the stall recovery procedure.
S3. Loss of pilot situational awareness leading to LOC-I	Task: Change altitude while maintaining airspeed in IMC.
S4. Energy management leading to performance decrement	Task: Change altitude while accelerating.

Subjective situational awareness was measured immediately after each flight scenario using 10-D Situational Awareness Rating Scale (SART; Taylor, 1990) as derived from a multi-dimensional characterisation of situational awareness consistent with the theory of perception, attention, and cognition (Endsley, 1995; Taylor, 1990). Participants, based on their task performance, subjectively rated each dimension on a seven-point scale. 3-D SART dimensions, attentional demand, attentional supply, and understanding factors were formed from 10-D version of the SART scale (Taylor, 1990). The overall SART score was calculated for each participant for each flight scenario. Pilot rating of task-related stress, as an indirect measure of psychological fidelity used a 24-item Short Stress State Questionnaire (Helton, 2004), as a rating of distress, engagement, and worry states. Stress state was measured pre-test and after each flight scenario and scored on a 5-point Likert scale

(Helton & Nöswall, 2010). In order to analyse the change in the stress profile, the post/pre ratio scores (i.e., differential state changes) (Helton & Nöswall, 2015) were calculated. The concept of presence, a human awareness phenomenon, was operationalised as a measure of attention effectiveness in virtual environments (Witmer & Singer, 1998). The 29-item Presence Questionnaire version 3 (PQ; Witmer et al., 2005) was administered post-test and measured presence through four dimensions (factors): involvement, sensors fidelity, adaptation and immersion and interface quality (Witmer et al., 2005). The overall composite PQ score was calculated as a sum of each factor for each participant. Pilot's wellbeing and acceptance were measured through standardized questionnaires administered post-test. Usability was operationalised as an indirect measure of pilot's acceptance (Burney et al., 2017; Golden et al., 2004; Holden & Rada, 2011). As recommended by Lewis and Sauro (2017), participant's estimation of system level usability was measured using a unidimensional, 10-item SUS questionnaire (Brooke, 1996). Cybersickness was operationalised as an indirect measure of pilot's wellbeing. Simulator Sickness Questionnaire (SSQ) was used to determine the severity of the sickness symptoms induced by the VR simulator (Kennedy et al., 1993) and experienced by participants. Three factors measure, i.e., nausea, disorientation, and oculomotor distress structure, as well as an overall sickness severity were analysed (Kennedy et al., 1993).

Results

This exploratory study examined the effects of pilot psychological experiences, i.e., task-related stress, presence, and cybersickness, on situational awareness during VR exposure during upset prevention and recovery training. SPSS statistical software was used in the data analysis. All results were considered significant at an alpha level $p = 0.05$.

Situational awareness

SART overall score is a function of attentional demand, attentional supply and understanding dimensions (i.e., 3-D), similarly scores for attentional supply and understanding are functions of a wider set of dimensions (10-D).

Overall SART score comparison. As the analysed data failed Mauchly's test of sphericity ($p < 0.05$), to evaluate the effects of the flight upset scenarios on participant's perceived situational awareness, a one-way within-subjects repeated ANOVA with a Greenhouse-Geisser correction was conducted for SART overall score. The different flight scenarios had significant effect on overall SART score, $F(1.88, 18.76) = 6.25$, $p < 0.05$, η^2 (partial) = 0.39, with medium magnitude of the effect (Cohen, 1992). Post-hoc pairwise comparisons with Bonferroni correction showed two significant differences between S1 VFR upset scenario and S3 IFR upset scenario ($M\Delta = 4.64$, $SE = 1.22$, $p = 0.02$) and between S2 stall scenario and S3 IFR upset scenario ($M\Delta = 7.27$, $SE = 1.44$, $p < 0.05$). No significant interactions were found between other flight scenarios.

SART Attentional demand score comparison. The different flight scenarios had significant effect on attentional demand score, $F(3,30) = 14.72$, $p < 0.001$, η^2 (partial) = 0.60, with large magnitude of the effect (Cohen, 1992). Post-hoc pairwise comparisons with Bonferroni correction showed three significant differences between S1 VFR upset scenario and S3 IFR upset scenario ($M\Delta = -7.09$, $SE = 1.02$, $p < 0.001$), between S2 stall scenario and S3 IFR upset scenario ($M\Delta = -7.91$, $SE = 1.52$, $p = 0.002$) and between S3 IFR upset scenario and S4 energy management ($M\Delta = 6.64$, $SE = 1.66$, $p = 0.02$). No significant interactions were found between other flight scenarios.

SART Attentional supply score comparison. The assumption of normality for SART Attentional supply score was verified by inspection of the normal Q-Q Plots and Shapiro-Wilk's test and it was determined to be normally distributed ($p > 0.05$) for most flight scenarios and marginally acceptable for second flight scenario ($p = 0.03$). Attentional supply score between different flight scenarios were not statistically different $F(3,30) = 1.86$, $p = 0.16$, η^2 (partial) = 0.16, with small magnitude of the effect (Cohen, 1992).

SART Understanding score comparison. Understanding score between different flight scenarios were not statistically different $F(3,30) = 2.59$, $p = 0.07$, η^2 (partial) = 0.21 with small magnitude of the effect (Cohen, 1992).

Task-related stress state

An effect significance for each flight scenario was followed up, where applicable, with post-hoc pairwise comparisons with Bonferroni correction to control for the probability of committing a type I error.

Engagement state task comparison. As the collected data failed Mauchly's test of sphericity ($p < 0.05$), to evaluate the effects of the VR exposure on participant's stress state, a one-way within-subjects repeated ANOVA with a Greenhouse-Geisser correction was conducted for SSSQ engagement dimension, for pre-test and each post-task reported stress level. A single factor, the flight scenario, was used during the analysis (i.e., VFR upset vs stall vs IFR upset vs energy management). It was determined that the effect of stress state change in engagement levels were not statistically different between pre-test and post-flight upset scenarios ($F(2.17, 21.86) = 0.81$, $p = 0.47$, η^2 (partial) = 0.08) with negligible magnitude of the effect (Cohen, 1992).

Distress state task comparison. It was determined that the effect of stress state change in distress levels did not elicit a statistically significant change between pre-test and post-S1 VFR upset scenario ($Z = -1.23$, $p = 0.22$), pre-test and post-S2 stall scenario ($Z = -1.08$, $p = 0.28$), pre-test and post-S3 IFR upset scenario ($Z = -0.32$, $p = 0.71$), or pre-test and post-S4 energy management scenario ($Z = -0.37$, $p = 0.72$).

Worry state task comparison. As the collected data failed Mauchly's test of sphericity ($p < 0.5$), to evaluate the effects of the VR exposure on participant's stress state, a one-way within-subjects repeated ANOVA with a Greenhouse-Geisser correction was conducted for SSSQ worry dimension, for pre-test and each post-task reported stress level. A single factor, the flight scenario, was used during the analysis (i.e., VFR upset vs stall vs IFR upset vs energy management). It was determined that the effect of stress state change in worry levels was not statistically different between pre-test and post-flight upset scenarios ($F(1.85, 18.50) = 2.04$, $p = 0.16$, η^2 (partial) = 0.17) with a weak magnitude of the effect (Cohen, 1992).

Presence

Considering the magnitude of scores for each PQ factor, the sensory fidelity ranked the lowest at 58.45% of the maximum score, followed by the involvement factor at 74.2%, and interface quality at 78.80%. Adaptation and immersion PQ factor ranked the highest at 85.5% of the maximum score. The overall results demonstrate a moderate presence score of 74.56%.

Pilot's wellbeing and acceptance

Cybersickness. The descriptive statistics associated with cybersickness as measured post-test using SSQ (Kennedy et al., 1993) are reported in Table 2. Considering the magnitude of scores disorientation ranked the highest most severe factor with the broadest range. The fullness head, blurred vision and dizziness were the symptom profiles elicited by the test conditions that impacted the most. Oculomotor disturbance being the second most severe factor was affected primarily by general discomfort, headache, eye strain and blurred vision symptom profiles. Nausea, the lowest symptomatic factor, was primarily impacted by general discomfort, increased salivation, and sweating. The overall results demonstrate moderate sickness severity, with a relatively broad range.

Threshold values originally proposed by Stanney (1997), would classify the severity as concerning, and extended sickness severity analysis shall be considered.

Table 2: Descriptive statistics for nausea, oculomotor, and disorientation cybersickness factors as measured using SSQ. Note: M = mean; N = frequency; SD = standard deviation.

SSQ Factor group	N	M	SD	Range
Nausea	11	10.47	11.64	0 ÷ 28.62
Oculomotor Disturbance	11	14.47	12.43	0 ÷ 37.90
Disorientation	11	18.98	25.87	0 ÷ 69.60
Overall sickness severity	11	16.32	16.23	0 ÷ 44.88

No participant dropout was recorded during the study. Singular VR exposure mean time was less than 10 minutes for all participants, meaning that each participant was exposed to the VR system for 50 minutes or less for the total duration. The exposure duration was not recorded as a variable but was measured during each test to control for possible exceedances.

System Usability. The SUS results demonstrate a high usability score (M = 86.13, SD = 8.90, Range = 70.00 ÷ 97.50) and it is graded as "excellent" according to SUS adjective rating (Bangor et al., 2008). Systems with scores of less than 72 as marginally acceptable and extended usability analysis shall be considered. Based on the positive score of the investigated immersive flight training system across four different flight upset scenarios, there is an indication that the system has a high acceptance rate among participants in the context of UPRT application.

Relationship analysis

This exploratory research aimed to determine the effects of the pilot's presence, task-related stress and cybersickness on situational awareness observed under four different flying upset scenarios. Correlational analysis was applied to determine the relationship between the variables. A nonparametric measure of Spearman's rank-order correlation was run.

It was determined that the strength of the relationship between situational awareness as measured with SART and self-reported stress state, as measured with SSSQ, was insignificant for the majority of scenarios. However, in-flight scenario three, there was a strong, positive correlation between SART attentional demand factor and SSSQ engagement factor, which was statistically significant ($r_s(9) = 0.61$, $p = 0.048$). Moreover, in-flight scenario four, a strong, negative correlation between SART overall score and SSSQ engagement factor, was also statistically significant ($r_s(9) = -0.75$, $p = 0.008$), as was for SART attentional supply and SSSQ engagement ($r_s(9) = -0.61$, $p = 0.049$), and between SART understanding and SSSQ distress ($r_s(9) = 0.62$, $p = 0.043$). It was determined that the strength of the relationship between situational awareness and self-reported presence, measured with PQ, was insignificant for all scenarios. Similarly, no significant relationship between situational awareness and cybersickness, measured with SSQ, was recorded. Negative correlation between cybersickness and another VR construct – presence, was recorded. The resulting strong negative correlation coefficient, $r_s (9) = -0.62$, was statistically significant, $p = 0.40$. Furthermore, a significant negative relationship was also found between presence and SSQ nausea factor ($r_s (9) = -0.79$, $p < 0.01$) and between presence and SSQ oculomotor disturbance factor ($r_s (9) = -0.63$, $p < 0.05$). The results from correlational analysis indicate potential interactions between the minority of variables. A small sample size should be considered as a partial explanation of lack of statistical significance of relationship between the variables of interest.

Discussion

This exploratory study aimed to establish whether a VR-enabled flight training device will provide effective UPRT training to ensure equivalent safety with the FSTD considering pilot situational awareness competency. It was achieved by determining the effects of the pilot's presence, task-related stress and cybersickness on situational awareness during upset prevention, and, if necessary, recovery and by assessing pilot acceptance of VR-enabled flight training device in UPRT. For this, professional and qualified pilots with sizeable flying experience were selected for the research. The results can be interpreted and implemented in other domains considering certain limitations of the research.

Application of situational awareness in prevention and manual control for recovery phases was the key criterium in the selection of the evaluative scenarios. Considering the overall SART score alone, two significant positive differences in self-reported level of situational awareness between VFR upset scenario (Scenario 1) and IFR upset scenario (Scenario 3) and between the stall scenario (Scenario 2) and IFR upset scenario (Scenario 3) were reported. This is expected and can indicate a lower overall level of situational awareness related to deprivation of visual cues in IMC conditions but increased demand for information exclusively from the instruments to substitute that constraint. Recognition of all required stimuli and processing the information in upset conditions are essential skills acquired by pilots during UPRT. Low attentional demand across VFR scenarios can suggest a low level of perceived variability, and complexity of the situation. High attentional supply across all scenarios indicated pilot's optimal arousal level, spare mental capacity, and good concentration. A moderate level of reported understanding dimension can indicate acceptable quality and quantity of information and general familiarity with the glass cockpit instruments used in the study. Considering perceived situational awareness alone, given the low level of prior exposure to the VR technology, and high proficiency in upset management (as confirmed during flight scenarios), the result of the analysis indicates a lack of negative consequences of VR application on situational awareness needed for effective upset prevention. One limiting factor related to the methodology employed to collect the data, regardless of the validity of the construct, the temporal characteristic of pilot's situational awareness must be considered to establish the full extent of the effect of the VR on situational awareness across the whole duration of the flight. Pilot situational awareness can be impaired when exposed to physical or psychological stress. As these symptoms affect predominantly perception it was difficult to interpret the results due to the adaptation of three factors, engagement, distress, and worry. The result indicates high engagement and low levels of reported distress and worry across all flight scenarios. High attentional demand observed in IMC flight conditions correlated with high engagement may suggest more cognitive resources diverted to perceive the information presented on flight instruments. Lower attentional supply resulted in a higher level of engagement, as observed in scenario four (energy management) would be expected as higher concentration, and attention is required limiting spare mental capacity and engaging the pilot more. Similar observation has been made between understanding factor and distress, as negative correlation demonstrates that the better information and familiarity with the VE, the lower distress. Although SSSQ method is claimed as valid and reliable (Helton, 2004), an objective methodology could be employed to factor in the dynamic character of the scenario and individual characteristics of pilots when performing under stress conditions (i.e., HRV or Galvanic Skin Response meters). It would be expected that stress affects pilot situational awareness during the upset, but the extent of the impact cannot be established. Furthermore, the use of VR in the study additionally complicates the interpretation of results and potentially dilutes content validity.

Presence scores were not found to be significantly correlated with situational awareness in all flight scenarios. High adaptation and immersion PQ factors correlate with high usability scores indicating low entry barriers and overall high engagement in the VR content. A low sensory fidelity score was expected as the interaction with the virtual content, lacks a naturalistic feeling of controlling virtual content and shall be considered primitive. Low sensory fidelity could potentially have an indirect effect on pilot's perception, but at this stage, this argument is not verified, and it is speculative. The results cannot support the argument that the construct of presence is affected by situational factors or vice-versa (He et al., 2018; Jung et al., n.d.; Prothero et al., 2016).

The results of the cybersickness survey demonstrate that the VR-based FTD delivers moderate sickness severity. Although the score of 16.2 is lower than other cybersickness severity scores published in the research literature (Kolasinski & Gilson, 1998; Saredakis et al., 2020; Stanney et al., 1997; Webb, 2010), caution needs to be exercised. The total severity score is close to being classified as concerning, and extended sickness severity analysis shall be considered. The results above 20 would automatically classify the system as a “bad simulator”, but as reported by Stanney (1997) the average sickness severity experienced by participants in virtual environment systems is on average three times higher than when using flight training devices. As expected, (Seay et al., 2002; Weech et al., 2019), a negative significant correlation between cybersickness and presence was recorded. The most probable cause is the diversion of participant attention form from unwanted symptoms, e.g., sensory conflict while experiencing higher presence levels.

The second aim of the study was to assess participant acceptance of VR technology-enabled flight training devices in UPRT. The concept of applying usability measures as a proxy of user acceptance (Burney et al., 2017; Golden et al., 2004; Holden & Rada, 2011) requires understanding of participant behaviour, and attitudes toward technology to guarantee effective, efficient, and satisfactory operation (Holden & Rada, 2011). A high usability score was recorded for the study, and the acceptance of the VR FTD was rated as “excellent” according to SUS adjective rating (Bangor et al., 2008) confirming positive and unsolicited feedback from participants during and after the study.

The key limitation of this study is exclusive employment of a sample with military aviation background with a sizeable flying experience. All participants indicated prior completion of UPRT or comparable training programme with experience in stall conditions and extreme attitudes, exceeding any UPRT requirements for pilots in the commercial aviation sector. While recognising wider upset training syllabubs, the core pilot competencies, as defined by ICAO, must be acknowledged if generalising the study.

The growing evidence suggests that upset-recovery training can be delivered using alternative VR-based TFD (Groen et al., 2012; Leland et al., 2009; Ommerli, 2019; Rogers et al., 2009). The results from this study partially confirm this claim. No evidence has been found that situational awareness was negatively affected by exposure to VR, with certain reactions to stimuli degradation (i.e., flight upset in IMC) resulting in a predictable outcome of increased attentional demand. Limited correlation has been found between situational awareness and task-induced stress, and no relationship was found between situational awareness and presence or cybersickness. The cybersickness severity score is close to being classified as concerning, and extended sickness severity analysis shall be considered. A high participant acceptance score was recorded for the study, and it is graded as “excellent” according to the adjective usability rating.

References

Airbus. (2022). Airbus Statistical Analysis of Commercial Aviation Accidents Statistical Analysis of Commercial Aviation Accidents.

Bangor, A., Kortum, P. T., & Miller, J. T. (2008). An Empirical Evaluation of the System Usability Scale. International Journal of Human–Computer Interaction , 24(6), 574–594. https://doi.org/10.1080/10447310802205776

BEA. (2012). Final Report on the accident on 1st June 2009 to the Airbus A330-203 registered F-GZCP operated by Air France flight AF 447 Rio de Janeiro - Paris.

Boeing. (2021). Boeing Statistical Summary of Commercial Jet Airplane Accidents - Worldwide Operations 1959-2020.

Brooke, J. (1996). SUS - a quick and dirty usability scale. In Usability Evaluation in Industry (1st ed.). Taylor & Francis.

Brooks, R., & Ransbury, P. (2019, August 16). Why Upset Prevention Training Alone Is Not Enough. Aviation Performance Solutions. https://apstraining.com/resource/why-upset-prevention-training-alone-is-not-enough/

Burney, S., Ali, S. A., Ejaz, A., & Siddiqui, F. A. (2017). Discovering the Correlation between Technology Acceptance Model and Usability. IJCSNS International Journal of Computer Science and Network Security, 17(11).

Cohen, J. (1992). A Power Primer. Psychological Bulletin [PsycARTICLES, 112, 155.

Endsley, M. (1995). Toward a theory of situation awareness in dynamic systems. Human Factors, 37(1), 32–64. https://doi.org/10.1518/001872095779049543

Endsley, M., & Jones, D. (2016). Designing for Situation Awareness An Approach to User-Centered Design. In Designing for Situation Awareness (2nd ed.). CRC Press. https://doi.org/10.1201/B11371

Endsley, M., Sollenberger, R. L., Nakata, A., & Stein, E. S. (2000). Situation Awareness in Air Traffic Control: Enhanced Displays for Advanced Operations.

Golden, W., Acton, T., Gudea, S., & Scott, M. (2004). Usability and Acceptance in Small-Screen Information Systems. ECollECTeR Conference Proceedings, 6.

Groen, E., Ledegang, W., Field, J., Smaili, H., Roza, M., Fucke, L., Nooij, S., Goman, M., Mayrhofer, M., Zaichik, L., Grigoryev, M., & Biryukov, V. (2012). SUPRA - Enhanced upset recovery simulation. AIAA Modeling and Simulation Technologies Conference 2012, 14. https://doi.org/10.2514/6.2012-4630

He, Z., Zhu, F., Perlin, K., & Ma, X. (2018). Manifest the Invisible: Design for Situational Awareness of Physical Environments in Virtual Reality. https://doi.org/10.48550/arxiv.1809.05837

Helton, W. S. (2004). Validation of a Short Stress State Questionnaire. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 48(11), 1238–1242. https://doi.org/10.1177/154193120404801107

Helton, W. S., & Nöswall, K. (2015). Short Stress State Questionnaire: Factor structure and state change assessment. European Journal of Psychological Assessment, 31(1), 20–30. https://doi.org/10.1027/1015-5759/A000200

Hockey, G. (1986). Changes in operator efficiency as a function of environmental stress, fatigue, and circadian rhythms. Handbook of Perception and Human Performance, 2, 44-1-44–49.

Holden, H., & Rada, R. (2011). Understanding the Influence of Perceived Usability and Technology Self-Efficacy on Teachers Technology Acceptance. Journal of Research on Technology in Education, 43(4), 343–367. https://doi.org/10.1080/15391523.2011.10782576

ICAO. (2014a). Doc 10011, Manual on Aeroplane Upset Prevention and Recovery Training. In Doc 10011 AN/506 (Issue 1, p. 102). International Civil Aviation Organization.

ICAO. (2014b). Manual on Aeroplane Upset Prevention and Recovery Training (Doc 10011 AN/506; p. 102). International Civil Aviation Organization.

ICAO. (2015). Doc 9625 Manual of Criteria for the Qualification of Flight Simulation Training Devices - Volume I - Aeroplanes. In DOC 9625-1 (Issue 4, p. 616). International Civil Aviation Organization.

Jones, D., & Endsley, M. (1996). Sources of situation awareness errors in aviation. Aviation, Space, and Environmental Medicine, 67(6), 507–512.

Jung, D., Jo, S., & Myung, R. (n.d.). A Study of Relationships between Situation Awareness and Presence that Affect Performance on a Handheld Game Console. Proceedings of the 2008

International Conference in Advances on Computer Entertainment Technology - ACE ’08. https://doi.org/10.1145/1501750

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator Sickness Questionnaire: An Enhanced Method for Quantifying Simulator Sickness. The International Journal of Aviation Psychology, 3(3), 203–220. https://doi.org/10.1207/S15327108IJAP0303_3

Kolasinski, E. M., & Gilson, R. D. (1998). Simulator Sickness and Related Findings in a Virtual Environment. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 42(21), 1511–1515. https://doi.org/10.1177/154193129804202110

Leland, R., Rogers, R. O., Boquet, A., & Glaser, S. (2009). An Experiment to Evaluate Transfer of Upset-Recovery Training Conducted Using Two Different Flight Simulation Devices.

Lewis, J. R., & Sauro, J. (2017). Revisiting the Factor Structure of the System Usability Scale. Journal of User Experience, 12(4), 183–192.

Miglior, M. (2014). UPRT and Simulators (p. 21). Ecair Aviation.

NTSB. (2009). Loss of Control on Approach Colgan Air Inc. Operating as Continental Connection Flight 3407 Bombardier DHC-8-400 N200WQ Clarence Center New York February 12 2009.

Ommerli, C. (2019). Examining the Effects of Perceived Telepresence, Interactivity, and Immersion on Pilot Situation Awareness During a Virtual Reality Flight Exercise [Carleton University]. https://doi.org/10.22215/etd/2020-13923

Prothero, J. D., Hoffman, H. G., Parker, D. E., Furness, T. A., & Wells, M. J. (2016). Foreground/Background Manipulations Affect Presence: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2, 1410–1414. https://doi.org/10.1177/154193129503902111

Richards, N. D., Gandhi, N., & Bateman, A. J. (2012). Improved upset recovery strategies through explicit consideration of pilot dynamic behavior. AIAA Guidance, Navigation, and Control Conference 2012, 15. https://doi.org/10.2514/6.2012-4821

Rogers, R. O., Boquet, A., Howell, C., & Dejohn, C. (2009). An Experiment to Evaluate Transfer of Low-Cost Simulator-Based Upset-Recovery Training. United States. Department of Transportation. Federal Aviation Administration. https://doi.org/10.21949/1503647

Saredakis, D., Szpak, A., Birckhead, B., Keage, H. A. D., Rizzo, A., & Loetscher, T. (2020). Factors Associated With Virtual Reality Sickness in Head-Mounted Displays: A Systematic Review and Meta-Analysis. Frontiers in Human Neuroscience, 14. https://doi.org/10.3389/FNHUM.2020.00096

Seay, A. F., Krum, D. M., Hodges, L., & Ribarsky, W. (2002). Simulator sickness and presence in a high field-of-view virtual environment. Conference on Human Factors in Computing Systems - Proceedings, 784–785. https://doi.org/10.1145/506443.506596

Slater, M., & 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. https://doi.org/10.1162/PRES.1997.6.6.603

Stanney, K. M., Kennedy, R. S., & Drexler, J. M. (1997). Cybersickness is Not Simulator Sickness. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 41(2), 1138–1142. https://doi.org/10.1177/107118139704100292

Stark, J. M., Comstock, J. R., Prinzel, L. J., Burdette, D. W., & Scerbo, M. W. (2001). A Preliminary Examination of Situation Awareness and Pilot Performance in a Synthetic Vision Environment. Proceedings of the Human Factors and Ergonomics Society, 40–43. https://doi.org/10.1177/154193120104500208

Steuer, J., & Reeves, B. (1992). Defining Virtual Reality: Dimensions Determining Telepresence. Journal of Communication, 42(4), 73–93.

Taylor, R. M. (1990). Situational awareness rating technique (SART): The development of a tool for aircrew systems design. In Situational Awareness in Aerospace Operations (pp. 111–127). NATO-AGARD-CP-478.

van de Merwe, K., van Dijk, H., & Zon, R. (2012). Eye Movements as an Indicator of Situation Awareness in a Flight Simulator Experiment. The International Journal of Aviation Psychology, 22(1), 78–95. https://doi.org/10.1080/10508414.2012.635129

Webb, C. (2010). Simulator Sickness in the MH-47G Simulator.

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and cybersickness in virtual reality are negatively related: A review. Frontiers in Psychology, 10(FEB), 158. https://doi.org/10.3389/FPSYG.2019.00158/XML/NLM

Witmer, B. G., Jerome, C. J., & Singer, M. J. (2005). The Factor Structure of the Presence Questionnaire. Presence: Teleoperators and Virtual Environments, 14(3), 298–312. https://doi.org/10.1162/105474605323384654

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