5-1-2021

Pilot Situation Awareness and Risk of Critical Incidents Using a Novel Online Flight Simulation Tool

Kirsten Brightman

Kathleen Van Benthem

Chris M. Herdman

PILOT SITUATION AWARENESS AND RISK OF CRITICAL INCIDENTS USING A NOVEL ONLINE FLIGHT SIMULATION TOOL

Kirsten Brightman[1]
Kathleen Van Benthem[1]
Chris M. Herdman[1]
1. Carleton University
Ottawa, Canada

Pilot situation awareness (SA) is a critical influence on decision making and an important element in maintaining the safe control of an aircraft. The present research investigated whether measures of pilot SA, gathered via an online computer-based cognitive screening tool for pilots, could be used to predict pilot's likelihood of real-world critical incidents. A risk score for each pilot was developed based on their self-reported critical incidents from their actual flight history. It was hypothesized that individuals with lower SA scores would have higher risk scores. The impact of age and pilot experience were also considered, as these factors are known to influence achievement of SA. Results report on comprehensive models of flight performance that quantify the effects of three levels of SA on risk.

Accident and fatality rates in general aviation (GA) have remained consistently high (Kenny, 2020) despite targeted safety strategies. Analysis of general aviation incidents indicates that approximately 70% of accidents are due to pilot error (Kenny, 2020). Specifically, SA errors are frequently linked to pilot-related accidents in general aviation (Bolstad et al., 2010, Jones & Endsley, 1996). Successful performance and safety outcomes rely heavily on pilot SA. Screening pilots for declines or deficits in SA may represent a potential approach to mitigate aviation accident and fatality rates. A current gap in the aviation domain are ecologically valid screening tools that identify pilots who may be at risk for SA failures.

SA is a well-studied element of pilot cognition and is recognized as a fundamental component in performance and safety outcomes in aviation. SA is thought to be a salient causal factor in aviation accidents. Notably, of major air carrier incidents, 88% of those involving human error could be attributable to problems with SA (Endsley, 1995a). SA has been identified as the most significant human factor causation in commercial air transport accidents (Kharoufah et al., 2018). In simulated flight SA is a significant contributor to performance (Bolstad et al., 2010) and predictive of safety outcomes (Van Benthem & Herdman, 2020). Endsley (1995a, p. 36) describes SA as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future." The safe operation of an aircraft depends upon the accuracy and completeness of the pilot's SA. SA is thought to build on three hierarchical levels of cognitive processing (Endsley,1995a). The first level involves the pilot's perception of relevant situational elements, their status and their characteristics without interpretation to the larger picture. Approximately 76% of SA errors can be attributed to failures in Level I (Jones & Endsley, 1996). Level II SA involves comprehending the significance of situation elements to current goals and circumstances. At this level, a pilot develops a mental model of the environment, in which elements and events are related. The third and final level of SA involves using knowledge regarding the status and significance of elements to predict future states of the environment.

Aging affects multiple elements of cognition necessary for generating and maintaining SA, such as working memory and attention. (Finnigan et al., 2011). As such, SA is thought to be negatively affected by age-related cognitive decline (Bolstad & Hess, 200). Pilots are not immune to typical cognitive aging (Hardy & Parasuraman, 1997) and there is a growing body of literature which suggests age has a significant negative impact on pilot performance. Older GA pilots have been shown to perform worse than younger pilots in simulated flight (Taylor et al., 2007, Van Benthem & Herdman, 2016). Some of the negative impacts of aging might be alleviated by experience as it has the potential to extend a pilot's working memory capacity through increased automaticity of relevant skills. However, there have been mixed conclusions regarding if expertise can compensate for age-related decrements in performance.

The present research explored whether measures of SA from an online cognitive screening tool for aviators can predict self-reported critical incident risk. Structural equation modeling (SEM) was used to model the direct effects of pilot age and experience on their ability to achieve SA at the three levels, as well as the subsequent direct effects of SA on self-reported critical incident data. SA measures were based on Endsley's three-level characterization. A risk score for each pilot was developed based on their self-reported critical incidents from their actual flight history. Three main hypotheses were investigated. First, increased age was hypothesized to negatively affect SA at all levels. Second, increased experience level was hypothesized to positively affect SA at all levels. Finally, decreased SA at all levels was hypothesized to be associated with higher risk scores.

Method

Participants

The sample was composed of 65 pilots, with ages ranging from 18 to 80 years (M = 48.78, SD = 12.47). Admission criteria include being a licensed/permitted pilot and holding a Canadian medical certificate. Pilots had between 46 and 26,500 logged flight hours (M=3376.01, SD=5570.31), and held an active license for a range of 1 to 54 years (M=20.59, SD=12.10). Certification level ranged from Student to Airline Transport, but the majority were Private VFR with 1 or more additional ratings. All participants provided informed consent to participate in the study in accordance with the Carleton University Research Ethics Board

Materials

Participants completed the study online using a personal electronic device. During the flight exercise, participants watched five short videos from the view of the left pilot's seat of a Cessna 172 Skyhawk. The virtual flight videos were delivered by a screen display and included the interior/exterior of the Cessna 172 Skyhawk and exterior terrain details. Underneath the flight video, two sliders were displayed which signified either a flight instrument or a mental state (e.g. mental workload, SA).

Procedure

Participants completed a pre-flight questionnaire regarding their expertise, demographics, and history of critical incidents. A list of seventeen possible pilot-caused critical incidents was provided. Participants selected all of those which they had personally experienced while acting as pilot in command. Next, participants completed one practice flight leg followed by four test legs. Throughout each video, participants were instructed to monitor and adjust the two sliders so that they accurately reflected the indicated flight instrument or the pilot's self-rated mental state. As pilots were not able to control the simulated aircraft, the slider task (e.g., matching slider values to actual altitude) was included as an alternative visual motor task. After completing each leg, pilots were asked questions relating to Level 1 SA (details regarding other aircraft heard in the radio call messages, and instrument panel monitoring) and Level 2 SA (Ownship and Conflict Detection [derived from radio call messages]). Responses to the questions were used to generate the SA variables.

Conceptual Framework

The current research presents a conceptual framework for understanding SA and pilot risk. Age, certification level, SA at each level were incorporated in the proposed model to examine their influences on risk for critical incidents. Level 1 SA variables related to recall of static information. This information came either from radio calls, such as other aircraft call signs, type, location and intention (SA Others & Intention Others) or was derived from pilot's own instruments (Instrument Error). Level 2/3 SA variables involved awareness of dynamic information (Ownship and Conflict Detection). The number of hours flown was included in the model as a control variable. The main outcome variable of the model was risk score. A risk score was generated for each participant using their responses to the critical incident questionnaire. Each critical incident was assigned a grade from 1 to 5. Grading of critical incidents was primarily based on fatality rates associated with the incident, with 5 representing the highest risk of a fatality. Grades were established and assigned based on expert advice and accident data from the Transportation Safety Board of Canada (2018), International Civil Aviation Organization (2020) and Joseph T. Nall Report (Kenny, 2020). To generate the risk score, the corresponding grades for each critical incident selected by a participant were summed.

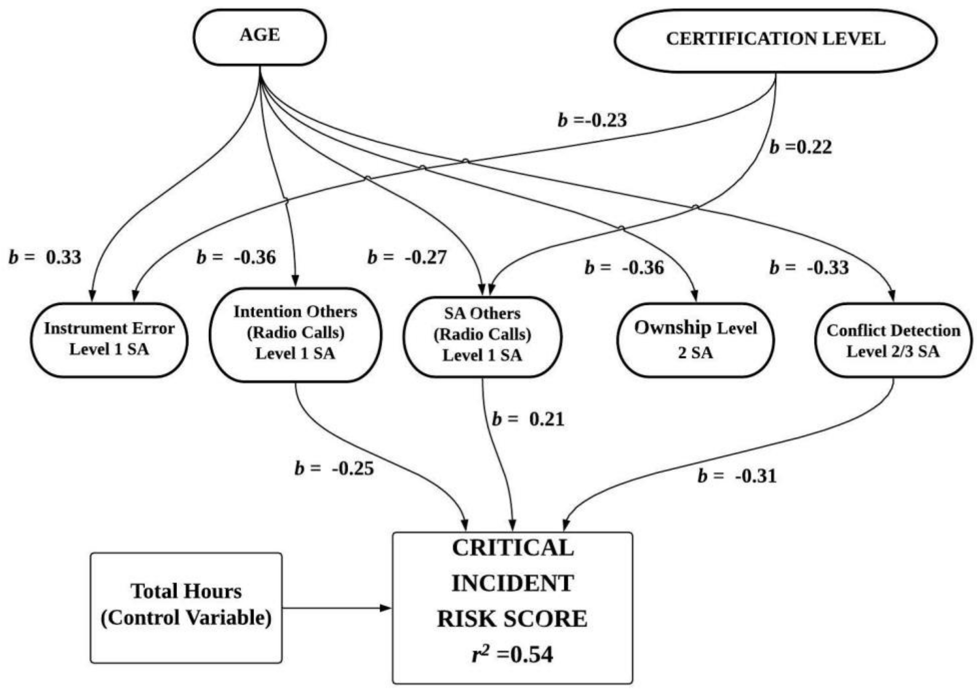
Results

Regarding the evaluation of the latent constructs (outer models), the final average variance extracted for all variables was above the 0.5 threshold, representing acceptable convergent validity. Using Cronbach's alpha and composite reliability as guides, all latent variables were found to have acceptable reliability (above 0.7). All final indicator loadings were above the threshold of 0.5 and considered acceptable. The inner model defines the relationship between the latent constructs and directly measures variables. All indices for assessing general fit and quality of the model were acceptable, and the structural model demonstrated a very good fit to the data. The r^2 associated with the outcome variable (Risk Score) and the path coefficients are essential measures for assessing the inner model. *r-Squared* is the standard method used to examine the predictive power of the structural model. It can be seen from Figure 1 that the model has a high predictive power, and accounts for 54% of the variance in the risk score. Figure 1 also demonstrates the path coefficients and *p*-values for each hypothesis, and it can be noted that most hypotheses were supported (non-significant paths were removed from the model).

As predicted, age significantly influenced all SA outcomes, such that increased age resulted in poorer SA scores. Certification level had a significant influence on Instrument Error (β= -0.228, p=.026) and SA Others (β= 0.223, p=.028) such that higher certification level resulted in better SA scores: this is consistent with the hypothesis that experience would positively affect SA. However, certification did not significantly influence Level 2 or 3 SA. Intention (β= -0.247, p=.017) and Conflict Detection (β= -0.308, p=.004) significantly affected Risk Score, such that individuals who performed better on these measures of SA had lower risk scores; partially supporting hypothesis 3. SA Others was found to significantly affect Risk Score (β= 0.208, p=.038), however the direction of the relationship was not as hypothesized. The results indicated an unexpected effect that those who performed better on the SA Others tasks also had higher risk scores. Instrument Error and Ownship were not significant influences on Risk score.

Figure 1.

Hypotheses Testing Results



Discussion

The present research aimed to model the direct and indirect effects of pilot age and experience on their ability to achieve SA at all three levels. In the present study age was shown to have a significant effect on all levels of SA. Overall as the age of the pilot increased their performance on SA tasks suffered. The finding that older age was associated with poorer SA outcomes is consistent with findings from the literature, which suggests that SA may be negatively affected by age-related cognitive decline. Age-related cognitive decline may affect Level 1 SA by limiting the amount of information which can be processed and the efficiency of retrieval processes (Bolstad & Hess, 2000). These effects may in turn diminish the development of accurate mental models (Level 2 SA) and future projections (Level 3 SA). The present research provides support for the account that cognitive-aging negatively affects pilot SA.

The effects of experience on the three levels of SA were also examined. As hypothesized, experience had a significant effect on Instrument Error and SA Others in the prediction direction.

The analysis showed there to be no significant relationship between experience and Ownship, Intention or Conflict. The finding that experience had significant effects on Level 1 SA but not Level 2 or Level 3 is consistent with results from Endsley et al. (2002) who reported that differences in SA between experience levels was most pronounced in Level 1.

The primary goal of the present study was to evaluate whether pilot's SA abilities are related to their history of critical incidents. The results of the SEM analysis partially supported the hypothesis that pilots who performed better on SA tasks would have lower Risk Scores. Of the five SA factors, three had significant effects on Risk Score; SA Others, Intention and Conflict. Intention and Conflict demonstrated directionality effects which were consistent with hypotheses, such that pilots who scored higher on these SA tasks had lower risk scores. These findings indicate that SA abilities could be a significant factor contributing to accidents. Of all SA factors, Conflict (Level 3 SA) was found to have the greatest influence on Risk Score. Mental projection is considered a demanding task which people struggle to perform well (Jones & Endsley, 1996). However, it seems that individuals who can perform this task accurately may have better safety outcomes. Jones and Endsley (1996) reported high numbers of SA incidents at the first and second level of SA. Individuals who can successfully generate and uphold Level 3 SA may have effectively avoided making common errors seen in Level 1 and 2. Based on our results, the ability to generate Level 3 SA is a significant and influential predictor of a pilot's history of critical incidents.

Unexpectedly, the direction of the relationship between SA Others and pilot risk showed that pilots who performed better on these SA tasks had higher Risk Scores. The opposite direction of effects for SA Others and pilot risk may represent a paradox with self-reporting critical incidents. For a pilot to report that they've experienced an incident involving a SA failure, the pilot has to be consciously aware that this failure has occurred. A pilot cannot report, for example, that they've had a loss of SA or a near-miss if they don't realize that this has happened. To identify that an SA incident occurred, the pilot has to be aware enough of themselves and their surroundings to recognize that there has been an error. Following this logic, we suggest that individuals with superior awareness may perform better on the SA tasks and also report more critical incidents.

The main purpose of this study was to model the relationship between SA, individual factors and pilot's history of real-life critical flights incidents. The implications of our results are applicable to the development of cognitive assessment tools in the aviation domain which may use SA abilities as a main predictor of safety. Virtual assessment tools which can accurately predict pilot risk have the potential to meaningfully improve safety outcomes in aviation. Knowledge of the precise cognitive processes underlying SA, individual differences predicting SA, and the influence of SA on end performance, will be of utmost importance in the development of cognitive assessment tools for aviators.

References

Bolstad, C. A., Endsley, M. R., Costello, A. M., & 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. https://doi.org/10.1080/10508414.2010.487013

Bolstad, C. A., & Hess, T. M. (2000). Situation awareness and aging. In *Situation awareness analysis and measurement* (277-302).

Endsley, M. (1995a). A taxonomy of situation awareness errors, human factors in aviation operations; *Proceedings of the 21st Conference of the European Association for Aviation Psychology (EAAP)*, *3*, 287–292.

Endsley, M. R. (1995b). Toward a Theory of Situation Awareness in Dynamic Systems. *Human Factors*, *37*(1), 32–64. https://doi.org/10.1518/001872095779049543

Finnigan, S., O'Connell, R. G., Cummins, T. D. R., Broughton, M., & Robertson, I. H. (2011). ERP measures indicate both attention and working memory encoding decrements in aging. *Psychophysiology*, *48*(5), 601–611. https://doi.org/10.1111/j.1469-8986.2010.01128.x

Hardy, D. J., & Parasuraman, R. (1997). Cognition and flight performance in older pilots. *Journal of Experimental Psychology: Applied*, *3*(4), 313–348. http://dx.doi.org.proxy.library.carleton.ca/10.1037/1076-898X.3.4.313

International Civil Aviation Organization. (2020). Safety Report 2020. Retrieved from: https://www.icao.int/safety/Documents/ICAO_SR_2020_final_web.pdf.

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

Kharoufah, H., Murray, J., Baxter, G., & Wild, G. (2018). A review of human factors causations in commercial air transport accidents and incidents: From to 2000–2016. *Progress in Aerospace Sciences*, *99*, 1–13. https://doi.org/10.1016/j.paerosci.2018.03.002

Kenny, D. J. (2020). *30th Joseph T. Nall report-general aviation accidents in 2018*. Frederick, MD: AOPA.

Taylor, J. L., Kennedy, Q., Noda, A., & Yesavage, J. A. (2007). Pilot age and expertise predict flight simulator performance. *Neurology*, *68*(9), 648–654. https://doi.org/10.1212/01.wnl.0000255943.10045.c0

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

Van Benthem, K., & Herdman, C. M. (2020). *A Virtual Reality Cognitive Health Screening Tool for Aviation: Managing Accident Risk for Older Pilots.* Manuscript submitted for publication.

Van Benthem, K., & Herdman, C. M. (2016). Cognitive factors mediate the relation between age and flight path maintenance in general aviation. *Aviation Psychology and Applied Human Factors*, *6*(2), 81–90. http://dx.doi.org.proxy.library.carleton.ca/10.1027/2192-0923/a000102