2013

Individual Pilot Factors Predict Simulated Runway Incursion Outcomes

Kathleen Van Benthem

Chris M. Herdman

INDIVIDUAL PILOT FACTORS PREDICT SIMULATED RUNWAY INCURSION OUTCOMES

Kathleen Van Benthem (kathy_vanbenthem@carleton.ca)
Institute of Cognitive Science, Carleton University, 1125 Colonel By Drive
Ottawa, Ontario, K1S 0A3 Canada

Chris M. Herdman (chris_herdman@carleton.ca)
Visualization and Simulation Centre, Carleton University, 1125 Colonel By Drive
Ottawa, Ontario, K1S 0A3 Canada

Runway incursions are a critical issue facing the aviation industry, with general aviation accounting for 77 percent of runway incursions involving pilot deviations. The present study reports on the value of individual pilot factors in predicting the outcome of a simulated surprise runway incursion. Significant predictors of runway incursion management were pilot rating, self-rated awareness of the impact of other aircraft on flight, and perceptions of the mental demands of the flight tasks. In light of the aviation industry's reliance upon pilots' self-monitoring of competence, strategies for reducing runway incursions can capitalize on evidence that higher self-rated ability to maintain a comprehensive mental model of relevant airspace, and better judgments of the mental demands of high workload conditions predicted superior incursion management outcomes.

Runway incursions and excursions are a critical safety issue facing the aviation industry, with costs estimated at a billion US dollars annually (Honeywell Aerospace, 2007). While reducing the occurrence of runway incursions is an obvious first step in dealing with this issue, knowing which pilots might be at risk when encountering an incursion upon final approach is another strategy for dealing with this aviation safety problem. One might think that the ability to deal with unexpected events is regularly assessed in general aviation; however, in accordance with a pilot's rating there is variation in the time between regulated evaluations of aviator competence. The two-year time lapse for private pilots between formal evaluations of competence means that self-monitoring of ability is a critical element in pilot safety. The present study provides a unique analysis of the individual pilot factors of age and experience, and pilot self-rating of situation awareness and mental workload as predictors of simulated unexpected runway incursion outcomes.

Individual Pilot Factors and General Aviation Flight Performance

To the authors' knowledge, no general aviation studies to date have investigated the association of pilot age, experience, and subjective self-ratings of situation awareness and workload with the outcome of a true "surprise" event. In order to explore the likely effects of individual pilot factors on an unexpected incursion event upon final approach, evidence pertaining to other general aviation tasks and individual pilot factors is reviewed.

Age

The issue of age and pilot performance has been studied in longitudinal and cross-sectional analyses from a large cohort of general aviation pilots with consistent findings that younger pilots tend to significantly outperform older pilots on key aviation tasks. (Kennedy, Taylor, Reade & Yesavage, 2010; Taylor, Kennedy, Noda, & Yesavage, 2007; Yesavage, Taylor, Mumenthaler, Noda, & O'Hara, 1999; Yesavage, Jo, Adamson, Kennedy, Noda, Hernandez, Zeitzer, Friedman, Fairchild, Scanlon, Murphy, & Taylor, 2011). Yesavage et al. (1999) reported on a sample of 100 pilots and found that overall simulated flight performance was reduced for the older pilots, with age predicting 22% of the overall variance in performance. In 2000 Taylor et al. reported that while four cognitive factors accounted for 45% of the variance in simulator scores (working memory/processing speed, visual memory, motor coordination, and tracking) age contributed significant variance in addition to the cognitive factors. Taylor et al. (2005; 2007) reported lower performance for older pilots when following air traffic control messages, traffic avoidance, cockpit instrument scanning, and approach and landing ability. In a small sample, low-fidelity simulator study of general aviation pilots, Coffey, Herdman, Brown and Wade (2007) found that older pilots missed significantly more rare critical events (near traffic and cockpit malfunctions) than younger pilots. Coffey et al. suggested that age-effects pertaining to "change detection" might be responsible for the older pilots missing more events both inside and outside the cockpit. In light of age-effects noted for other aviation tasks, some requiring quick decision-making and traffic avoidance, it is posited that pilot age might be a predictor of performance handling an unexpected runway incursion upon final approach.

Flight experience

Taylor et al. (2005) found that older pilots performed worse on a communication task and that higher levels of expertise did not offer protection to the older pilots regarding accuracy of communication performance. Taylor et al. (2007) later found that higher-rated pilots showed fewer declines over time, and expertise offered some protection against declines in communication task performance. Yesavage et al. (2011) also reported that higher rating was associated with better performance for all pilot age groups; however, higher rating did not demonstrate mediation of the age-effects found in this sample. Kennedy, Taylor, Reade & Yesavage (2010) demonstrated that expertise provided an advantage for older pilots on the flight control task of banking performance in a holding pattern. Visual flight rules flight into instrument meteorological conditions is a common pilot violation associated with crashes in general aviation and, as such, has been studied in aviation simulation studies. Causse, Dehais, Arexis, and Pastor (2011) also investigated how pilot experience influenced weather-related go-arounds upon approach and found that pilot experience was a significant predictor of weather-related decision-making. It is proposed, that higher levels of experience, as indexed by rating, will predict better runway incursion management.

Subjective Measures of Workload and Situation Awareness

Subjective Measures of Workload

A pilot's ability to subjectively and accurately rate cognitive workload should reflect an appropriate appreciation of the load caused by current task demands. It is suggested that a pilot who is sensitive to the effects of the task demands should also have a good appraisal of the tasks and the environment in which they take place. A commonly used subjective measure of workload is the NASA TLX Index (Hart and Staveland, 1988). In a review of mental workload in the aviation domain Cain (2007) reported that subjective rating scales have shown some reliability and validity, but that the correlation with objective measures of workload has not been consistent. Vidulich (1988) suggested that subjective measures of workload may be poor predictors of actual performance and describes a series of laboratory experiments where task workload manipulations were not appropriately reflected in subject reports of workload. The present study will explore the utility of subjective workload measures in predicting performance in naturalistic setting where alignment between manipulated and subjective workload might be augmented.

Subjective Measures of Situation Awareness

Situation awareness is best described as the pilot perceiving and integrating relevant stimuli within a meaningful "volume of time and space" and using selected stimuli to build a mental model of the environment and project that model into the future (Endsley, 1988, p. 97). Vidulich, Crabtree, and McCoy (1993) found that a one-item scale representing situation awareness did not adequately represent actual situation awareness performance. SART (Taylor, 1990) is a commonly used three-dimensional self-measure of situation awareness. However SART items, which include rating of task demands and attention supply, have been criticized for reflecting workload constructs rather than situation awareness (Jones, 2000). It is proposed that a scale with multiple items indexing situation awareness pertaining to other aircraft locations and relevancy might be associated with performance handling unexpected events that occur during flight.

Method

Participants

The present study was comprised of a total of 108 pilots. Fifteen pilots had incomplete data for the subjective rating scales and were removed from the current analysis. The remaining 93 pilots were aged 19 to 81 years (*M*=46.5) and held a current medical certification. Pilot expertise included student, private (no additional ratings), private (additional ratings), and advanced (airline transport, commercial, and military) ratings. Table 1 reports the mean age, total flight hours, and recent pilot-in-command hours licensed for each pilot-rating group.

Table 1.
Pilot Rating-Group Characteristics

Rating Group (*N*)	Mean Age (years)	Mean Total Flight Hours*	Mean Pilot-in-Command Hours (past 12 months)*
1. Student (12)	40.1	49.2	6.1
2. Private (no additional rating) (35)	48.8	242.0	17.0
3. Private (additional rating e.g. Instrument Flight Rules) (27)	48.8	681.3	29.7
4. Advanced (ATP, Commercial, Military) (18)	43.3	2038.0	82.2

Note. * denotes significant ANOVA between group differences p<.05.

Flight simulator environment

Pilots flew a Cessna 172 non-motion simulator with instruments and controls integrated with Microsoft® Flight Simulator X. Three large screens positioned in an arc in front of the cockpit provided approximately 120 degrees of horizontal and 45 degrees of vertical field of view. Feedback from experienced pilots and flight instructors indicated that the simulator was representative of a Cessna 172 type aircraft.

Procedure

Pilots completed a consent form and a demographic and experience questionnaire before receiving detailed instructions regarding the simulator tasks. After an orientation and practice phase pilots were required to fly three left-hand patterns in a low cognitive workload condition and then three left-hand patterns in a high cognitive workload condition. Only the high workload condition is described as it is the condition of interest in the present study. In the high workload condition the airfield was uncontrolled, the terrain was mountainous, and the pilot interacted with up to four other computer generated aircraft. Pilots were required to provide details of their call sign, aircraft type, and location at routine points during the circuit via radio communication, and note the similar information provided by the "pilots" from relevant computer generated aircraft. The unexpected runway incursion event always occurred at the end of the final pattern in the high workload condition. Several other measures of pilot performance were collected but are not reported on in this analysis.

Measures

Simulated runway incursion outcome measure. The runway incursion outcome measure comprised the last task in the flight simulation protocol. In the final high workload condition the pilot was requested to fly two *touch-and-go* patterns. While the pilot was on final approach a "rogue" simulated aircraft was introduced at the distal end of the runway and traveling in the wrong direction (i.e. towards the oncoming pilot participant). To successfully manage this runway incursion pilots must have accomplished three main tasks in quick succession. The initial requirement was to correctly perceive and interpret the environmental stimuli; the second task was to use the interpreted information in order to make a decision regarding the required procedure to avoid the quickly approaching offending aircraft; finally, the correct decision regarding evasive action in the air must be selected and acted upon in timely and efficient manner. The runway incursion outcome score was based on the quality, timing and result of pilot response to the offending aircraft and could range from 0 (did not notice the incursion) to 10 (noticed the incursion with adequate time to make a radio call and perform a safe maneuver to avoid the incursion with no dangerous loss of separation between aircraft): scores below 5 indicate high risk for a poor outcome.

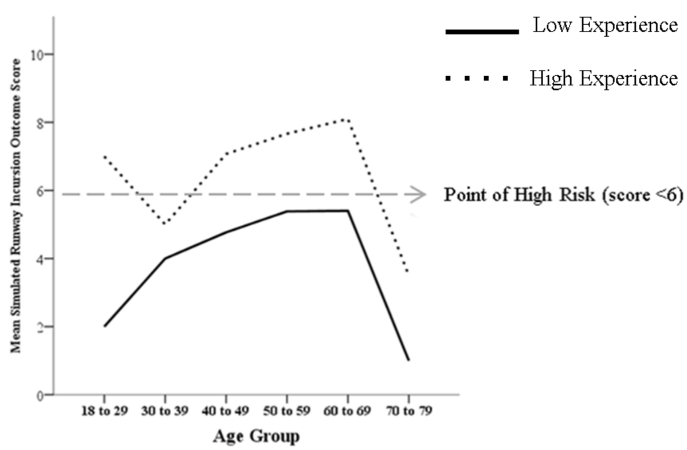
Pilot subjective ratings. A modified (for clarity of instructions and tasks) version of the NASA Task Load Index (TLX) (Hart and Staveland, 1988) was used to gauge pilots' perception of workload in both workload conditions. There were seven categories of workload on a continuous 100mm (end points low and high) rating scale: physical demand, mental demand, temporal demand, performance, effort, and frustration. Subjective situation awareness was measured using a 7-point scale designed specifically for aviation tasks at the ACE Laboratory at Carleton University and was based on the scale first reported by Hou, Kobierski and Brown (2007). The scale consisted of 11 items representing various aspects of situation awareness encountered during flight and included reference to awareness of the environment, tasks, time, priorities, relative position of ownship and other aircraft, future events, and overall level of situation awareness.

Results

Age, Experience, and Runway Incursion Outcome

Figure 1.

Mean Simulated Runway Incursion Outcome by Age and Experience Groups.



Note. Student and private pilot with no additional ratings are Low Experience pilots
Private pilots with additional ratings and ATP, commercial, and military are High Experience pilots

It is clear from Figure 1 that both younger and older pilots performed worse than the middle-aged pilots. Figure 1 suggests that age effects on runway incursion management might be obscured by the tendency of younger pilots to also score poorly on the incursion task. Not surprisingly, there was a marginally significant cubic relationship between age and incursion management, $F(3, 89)=2.43$, $p=.07$. Figure 1 also demonstrates the effect of experience on incursion management scores. Both low and high experience pilots improved their scores after age 30 and then saw a reduction in performance after age 70.

A one-way ANOVA showed a significant overall effect of pilot rating, $F(3, 89) = 9.05$, $p = .001$, $\eta_p^2 = .23$. Figure 1 suggests that highly experienced younger pilots might gain additional benefits for experience not enjoyed by the oldest pilots. However, the age x experience interaction was non-significant, $p > .05$.

Subjective Self-rating of Workload, Situation Awareness, and Runway Incursion Outcome

Pilots who recognized the higher level of mental demand in the high workload condition also tended to better manage the runway incursion. The mental demand item on the workload rating scale was the only item to correlate significantly with the runway incursion outcome.

Table 2.

Significant Predictors of Simulated Runway Incursion Outcome.

Predictor	Linear Regression Coefficients: Standardized Beta Coefficient	*t*- value
Pilot Rating*	.35	3.66**
SA of Impact of Other Aircraft on Flight*	.23	2.46*
TLX: Mental Demands of the Task	.21	2.25*

Note. * denotes significant predictor at p<.05
** denotes significant predictor at p<.01

Pilots who had higher self-ratings for situation awareness pertaining to how other aircraft in the system would impact their ability to fly the “perfect” pattern also showed better incursion management. A linear regression model revealed that together, pilot rating and self-rated mental workload and situation awareness accounted for 26% of the variance in runway incursion outcome, $F(3, 87) = 10.22$, $p < .001$. Two additional self-rated items of situation awareness showed marginal ($p < .1$) association with incursion management: awareness of task priorities and position of ownship relative to other aircraft. However, these items did not contribute significantly to the regression model.

Discussion

This study investigated whether pilot age, experience and pilot self-rating of situation awareness and workload predicted surprise simulated runway incursion outcome. With respect to age, the oldest group and the youngest pilot group with low ratings achieved mean scores that represent high risk (score less than 6/10) when encountering a runway incursion upon final approach. The marginally significant cubic relationship between age and runway incursion outcome scores supports this conclusion with larger samples perhaps defining this relationship further. Pilot rating contributed the most unique variance in runway incursion management scores. Subjective measures indexing a pilot’s self-rating for mental demands and comprehensive situation awareness factors pertaining to the interpretation of the impact of other aircraft on one’s own flight were also significant predictors of the outcome of the surprise incursion upon final approach.

Stakeholders addressing runway safety can capitalize on the evidence that strategies to reduce risk for pilots when encountering a runway incursion are best directed towards pilots with lower rating and poorer self-evaluation of situation awareness of other aircraft and poorer appreciation of the mental demands required in high workload conditions. A variety of methods to encourage targeted pilot self-evaluation between regulated biennial assessments of competence should be considered.

Acknowledgements

Funding for this research is provided in part by scholarships to K. Van Benthem from the Social Sciences and Humanities Research Council of Canada and the J. James Mackie Endowment for Graduate Scholarships, Ottawa, Canada. Research infrastructure was supported through a Canadian Foundation for Innovation and Ontario Innovation Trust grant to C. M. Herdman. Authors acknowledge Anne Barr and Andrew Staples of for their simulation engineering expertise.

References

Causse, M., Dehais F., Arexis, M. & Pastor, J. (2011). Cognitive aging and flight performances in general aviation pilots. Aging, Neuropsychology and Cognition, *18*(5) 544-561.

Coffey, E., Herdman, C.M. & Brown, M. (2007). Age-related changes in detecting unexpected air traffic and instrument malfunction. *Proceedings of the 14th International Symposium on Aviation Psychology, Columbus, OH.*

Endsley, M.R. (1988). Situation awareness global assessment technique (SAGAT). *Proceedings of the National Aerospace and Electronics Conference (NAECON)*, 789–795. New York: IEEE.

Endsley, M.R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M.R. Endsley & D.J. Garland (Eds.), *Situation awareness analysis and measurement*. Mahwah, NJ: LEA.

Hart, S.G. & Staveland, L.E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. Human Mental Workload. P.A.M. Hancock, N. Amsterdam, North-Holland: 139-183.

Honeywell Aerospace. (2009). *Honeywell smart runway and smart landing: Reducing the Risk of runway incursion and excursions*. Retrieved April 15, 2011 from www.honeywellrunwaysafety.com. © 2009 Honeywell International Inc.

Hou, M., Kobierski, R., & Brown, M. (2007). Intelligent adaptive interfaces for the control of multiple UAVs. *Journal of Cognitive Engineering and Decision Making, 1*(3), 27-28-362.

Jones, D. (2000). Subjective measures of situation awareness. In M.R. Endsley & D.J. Garland (Eds.), *Situation awareness analysis and measurement*. Mahwah, NJ: LEA.

Kennedy, Q., Taylor, J. L., Reade, G. &Yesavage, J. (2010). Age and expertise effects in aviation decision making and flight control in a flight simulator. *Aviation, Space, and Environmental Medicine, 81*, 489-497.

Taylor, R. M. (1990). Situational Awareness Rating Technique(SART): The development of a tool for aircrew systems design. *AGARD, Situational Awareness in Aerospace Operations 17 p(SEE N 90-28972 23-53).*

Taylor, J., O'Hara, R., Mumenthaler, M., Rosen, A. &Yesavage, J. (2005). Cognitive ability, expertise, and age differences in following age-traffic control instructions. *Psychology of Aging, 20*, 17–33.

Vidulich, M.A. (1988). The cognitive psychology of subjective mental workload. Human Mental Workload. P.A. Hancock and N. Meshkati. Amsterdam, NL, Elsevier Science Publishers B.V. (North-Holland): 219-229.

Vidulich, M. A., Crabtree, M. S., & McCoy, A. L. (1993). Developing subjective and objective metrics of pilot situation awareness. In Proceedings of the 7th International Symposium on Aviation Psychology (pp. 896-900).