Initial Cognitive Performance Predicts Longitudinal Aviator Performance

Jerome A. Yesavage,[1,2] Booil Jo,[1] Maheen M. Adamson,[1,2] Quinn Kennedy,[1] Art Noda,[1]
Beatriz Hernandez,[1] Jamie M. Zeitzer,[1,2] Leah F. Friedman,[1] Kaci Fairchild,[1,2] Blake K. Scanlon,[1,2]
Greer M. Murphy Jr.,[1] and Joy L. Taylor[1,2]

[1]Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, California.
[2]Department of Veterans Affairs Health Care System and Sierra-Pacific Mental Illness Research, Education, and Clinical Center, Palo Alto, California.

Objectives. The goal of the study was to improve prediction of longitudinal flight simulator performance by studying cognitive factors that may moderate the influence of chronological age.

Method. We examined age-related change in aviation performance in aircraft pilots in relation to baseline cognitive ability measures and aviation expertise. Participants were aircraft pilots ($N = 276$) aged 40–77.9. Flight simulator performance and cognition were tested yearly; there were an average of 4.3 ($\pm$ 2.7; range 1–13) data points per participant. Each participant was classified into one of the three levels of aviation expertise based on Federal Aviation Administration pilot proficiency ratings: least, moderate, or high expertise.

Results. Addition of measures of cognitive processing speed and executive function to a model of age-related change in aviation performance significantly improved the model. Processing speed and executive function performance interacted such that the slowest rate of decline in flight simulator performance was found in aviators with the highest scores on tests of these abilities. Expertise was beneficial to pilots across the age range studied; however, expertise did not show evidence of reducing the effect of age.

Discussion. These data suggest that longitudinal performance on an important real-world activity can be predicted by initial assessment of relevant cognitive abilities.

Key Words: Age—Aviation expertise—Cognitive performance—Executive function—Flight simulator performance—Processing speed.

PROCESSING-SPEED theory proposes that much of the age-related decline observed in measures of fluid cognition can be accounted for by age-related slowing in the speed of elementary cognitive operations (Salthouse, 1996). A large body of research shows that both behavioral and brain measures of processing speed can account for a substantial proportion of age-related variance in fluid cognitive abilities (e.g., Birren & Fisher, 1995; Bucur et al., 2008; Kwong See & Ryan, 1995; Lindenberger, Mayr, & Kliegl, 1993; Salthouse, 1991; Taylor, O'Hara, Mumenthaler, Rosen, & Yesavage, 2005; Walhovd et al., 2005). Recently, a large study of twins confirmed that genetic variance for processing speed was a leading indicator of variation in a number of age-related cognitive changes (Finkel, Reynolds, McArdle, Hamagami, & Pedersen, 2009).

Nonetheless, a processing speed model alone does not necessarily explain certain aspects of normal age-related decline in fluid abilities (Bugg, Zook, DeLosh, Davalos, & Davis, 2006; Schaie, 1989; Schretlen et al., 2000). In particular, decline in executive function has been proposed as a second mechanism underlying age-related decline in fluid cognition (Buckner, 2004; Dempster, 1992; Hasher & Zacks, 1988; Hasher, Zacks, & May, 1999; West, 1996).

Although the best way to measure executive function remains a challenge, it has been suggested that executive function be characterized as a collection of related cognitive control processes, in particular, "updating" working memory representations, "shifting" mental set, "inhibiting" automatic or prepotent response to irrelevant information, and "coordinating" complex, sequential behaviors to reach a goal (Friedman et al., 2008). A number of behavioral and neuroimaging studies have demonstrated links among executive control processes and age-related differences in performance of working memory and episodic memory tasks that stress interference, temporal processing, recollection, or contextual binding (Gray, Chabris, & Braver, 2003; Gunning-Dixon & Raz, 2003; Head, Kennedy, Rodrigue, & Raz, 2009; Hedden & Park, 2001; Kwong See & Ryan, 1995; Taylor et al., 2005). Together the results of these studies imply that when examining the explanatory power of processing speed, it may be advantageous to also examine other cognitive abilities, such as executive function.

Relatively few studies have examined the processing-speed theory from a longitudinal perspective (Hertzog, Dixon, Hultsch, & MacDonald, 2003; Schaie, 1989; Zimprich & Martin, 2002). A processing speed measure

collected at entry into a longitudinal study could predict in whom decline is more likely to occur. This prediction is based on the proposition that entry scores carry some rate-of-decline information. (That is, a person's score at study entry reflects the combination of the young-adult level plus an individual variation in the rate of decline, or growth, over time.) Lower processing speed scores at study entry may indicate risk for continued faster decline. Prior findings from the Religious Orders Study lend some support to the hypothesis that initial speed predicts subsequent cognitive decline (Wilson et al., 2002). First, substantial individual differences in rate of change were observed for several ability measures that involve fluid cognition (specifically, composite measures of "perceptual" speed, working memory, visuospatial ability, and memory retention). Second, initial level of performance on the speed measure correlated modestly with rate of decline in speed. Third, participants' rates of change on these ability measures were correlated (Wilson et al., 2002). Taken together, these findings suggest that assessing processing speed at study entry might carry enough rate-of-decline information to predict change in other measures that involve fluid cognitive abilities. Longitudinal work such as this provided the rationale and feasibility for our hypothesis that initial speed will predict longitudinal aviation performance.

Aviation is an ideal performance domain for examining the trade-off between age-related decline in basic cognitive abilities and the accumulation of experience. Concerns about age-related declines have favored mandatory age-based retirement of commercial airline pilots since 1960. The balance tipped slightly toward experience in 2007 when H.R. 4343–110th Congress: Fair Treatment for Experienced Pilots Act (2007) was signed into law, which increased the retirement age from 60 to 65 in the United States. The benefits of aviation expertise (variously defined) have been documented in many cross-sectional studies. These include laboratory studies of scanning cockpit instruments (Bellenkes, Wickens, & Kramer, 1997; Kasarskis, Stehwien, Hickox, Aretz, & Wickens, 2001), processing air-traffic control communications (Morrow, Leirer, Altieri, & Fitzsimmons, 1994; Morrow, Menard, Stine-Morrow, Teller, & Bryant, 2001; Morrow et al., 2003, 2005; Taylor et al., 2005), making aviation-related decisions (Doane, Sohn, & Jodlowski, 2004; Schriver, Morrow, Wickens, & Talleur, 2008; Wiggins & O'Hare, 1995), performing instrument flight maneuvers (Kennedy, Taylor, Reade, & Yesavage, 2010), and timesharing (Lassiter, Morrow, Hinson, Miller, & Hambrick, 1996; Tsang & Shaner, 1998). In studies that examined age differences, several found evidence that expertise moderated (reduced) age differences (Kennedy et al., 2010; Lassiter et al., 1996; Morrow et al., 1994, 2003; Tsang & Shaner, 1998), particularly when demands on memory encoding and storage were relatively low, or when vastly different levels of expertise were compared (e.g., pilots *vs.* nonpilots). All found that expertise benefited aviation performance.

For over 10 years, we have studied a cohort of aviators aged 40–70 years and more to better understand how their flight simulator performance changes as they approach and pass through their 60s (e.g., Taylor, Kennedy, Noda, & Yesavage, 2007; Yesavage, Taylor, Mumenthaler, Noda, & O'Hara, 1999). The purpose of the present work was to test the hypothesis that baseline cognitive processing speed and pilot expertise predict age-related change in aviator performance. However, there are other cognitive abilities such as executive function that are relevant for flight performance and decline with age (Morrow et al., 2001; Taylor, O'Hara, Mumenthaler, & Yesavage, 2000; Taylor et al., 2005). Because we have a rich data set that includes measures of other cognitive abilities, as an exploratory analysis, we used signal detection analysis (Kraemer, 1992) to identify additional cognitive predictors.

METHOD

Participants

Participants ($N = 276$) were part of the ongoing longitudinal Stanford/VA Aviation Study approved by the Stanford University Institutional Review Board. Enrollment criteria were age between 40 and 69 years, current Federal Aviation Administration (FAA) medical certificate, and total lifetime flying activity between 300 and 15,000 hours of flight time. All participants gave written informed consent to participate in the study, with the right to withdraw at any time. At entry, each participant was classified into one of the three levels of aviation expertise depending on which FAA pilot proficiency ratings they had previously attained: (a) least expertise: VFR (rated for flying under visual flight rules only); (b) moderate expertise: IFR (rated for instrument flight); and (c) high expertise: CFII and/or ATP (certified flight instructor of IFR students and/or certified to fly air-transport planes). Each rating requires progressively more advanced training and more hours of flight experience. All of the VFR pilots were recreational pilots, although a small minority was employed in aviation-related jobs such as aircraft sales or mechanics. The majority of the IFR group were recreational pilots, but approximately one tenth of the IFR group were certified flight instructors of VFR student pilots, aviation analysts, or had been aviators during military service. Approximately one half of the CFII/ATP participants were either air-transport pilots, CFIIs, or had job duties that included aircraft piloting. Table 1 lists demographic and flight expertise characteristics of the sample. Stepwise logistic regression modeling did not show indications of selective attrition by age, expertise, or cognitive measures at study entry (α level = 0.05). Pilots completed flight simulation and cognitive assessments yearly. There were an average of 4.3 data points per participant (± 2.7; range 1–13), spanning an average of 3.8 years (± 3.2; range 1–13).

Table 1. Participants' Demographic and Flight Experience Characteristics at Study Entry by Pilot Expertise Level

	Pilot expertise level			
	Least ($n = 70$)	Moderate ($n = 151$)	High ($n = 55$)	Total ($N = 276$)
Age in years				
M (SD)	56.7 (7.3)	58.5 (6.3)	55.5 (6.5)	57.5 (6.7)
Range	43–69	41–69	44–68	41–69
Years of education				
M (SD)	16.7 (2.2)	17.1 (1.9)	17.2 (1.9)	17.0 (2.0)
Women				
n (%)	12 (17.1)	22 (14.6)	4 (7.3)	38 (13.8)
White, non-Hispanic %	86	99	93	95
Total log hours				
M (SD)	980 (1,288)	1,876 (1,881)	5,260 (2,926)	2,323 (2,514)
Log hours in past month				
M (SD)	5.7 (7.1)	8.7 (10.1)	15.4 (17.3)	9.3 (11.7)

Note: Log hours are the flight hours pilots document in their log books, that is, a measure of experience.

Equipment

Pilots "flew" in a FAA-approved Frasca 141 flight simulator (Urbana, IL). Motion, vibration, and sound elements were not incorporated into the simulator protocol. This simulator was similar to a Cessna 172 fixed-wing aircraft. It was linked to a computer designed for graphics (Dell Precision Workstation and custom C++ OpenGL Linux software, Red Hat, Raleigh, NC) that generated a "through-the-window" visual environment and continuously collected data concerning the aircraft's position and communication frequencies. The simulator was located in a quiet darkened room kept at a comfortable temperature. The cockpit was illuminated by independent lighting separate from the light of the projector display. The visual display was projected on a screen 15 feet in front of the pilot, which allows a normal view of objects to be seen at a distance. Previous work in our laboratory supports the validity of the simulator as a method for measuring piloting performance because it has distinguished between novice and expert aviators, and between younger and older aviators (Taylor et al., 2007).

Measures

Flight simulator performance (dependent measure).— The scoring system of the flight simulator-computer system produces 23 variables that measure deviations from ideal positions or assigned values (e.g., altitude in feet, heading in degrees, airspeed in knots), or reaction time (in seconds). Because these individual variables have different units of measurement, the raw scores for each variable were converted to z-scores, using the initial visit mean and standard deviation (the scores of the two flights pilots flew each test day were averaged). The z-scores for the individual measures were aggregated on the basis of previous principal component analyses into four component measures: (a) accuracy of executing the air traffic control (ATC) communications regarding the heading, altitude, radio frequency, and transponder code (correct *vs* incorrect); (b) ability to avoid

traffic that appears 1200′ in front of the aircraft either 25′ above or below flight past and 25′ left or right of flight path where the pilot is expected to deviate to the opposing quadrant from the traffic as measured by distance from the traffic and correctness of the maneuver; (c) scanning cockpit instruments to detect engine emergencies for carburetor icing and oil pressure drop as measured by stopwatch seconds until observed; and (d) executing a visual approach to landing in terms of equally weighted left/right and above/below deviations from an ideal flight path. In addition, an overall flight performance summary score was computed as the mean of the four component scores.

Cognitive testing (predictor measures).—The cognitive predictors were seven composite scores derived from Cog-Screen-AE (Kay, 1995), a computer-administered battery of 13 subtests designed for screening and monitoring the perceptual and cognitive abilities relevant to flying. Over several years, factor structures have been proposed for the battery's 24 scores; the composite scores used in this study are a slight modification of those used in our prior work (Taylor et al., 2000). The composite scores for the baseline visit are presented in Table 2. Full descriptions of the subtests are available online (http://www.cogscreen.com/) and in the CogScreen-AE manual (Kay, 1995).

Procedure

Before the baseline visit, participants had six practice flights in the simulator to gain familiarity with the flight scenario used throughout the study. Participants typically completed their familiarization flights during a 1- to 3-week period, after which they had a 3-week break before returning for the baseline visit. At the baseline visit and each annual time point thereafter, the participant flew a 75-min flight in the morning and a 75-min flight in the afternoon. Each flight was followed by a 40- to 60-min battery of cognitive testing, including the CogScreen-AE. The entire test

Table 2. CogScreen Composite Scores Used in Exploratory Signal Detection Analysis

Composite score	n	M	SD	Minimum	Maximum
Processing speed	184	−0.02	0.37	−1.66	1.58
Executive function	174	−0.05	0.52	−2.08	1.35
Symbol-digit recall (% accuracy)	183	75.12	24.02	17.00	100.00
Working memory updating (% accuracy)	182	82.87	14.83	24.00	100.00
Working memory manipulation	184	0.01	0.47	−1.70	2.29
Motor coordination	184	0.40	0.90	−2.22	5.37
Tracking (error score)	184	−0.08	0.50	−1.95	2.45

day lasted approximately 6 hr, including a 30- to 50-min lunch break. The simulation was scheduled during normal working hours from 9:00 a.m. to 4:00 p.m. Test days were repeated annually.

Each flight began with ATC takeoff clearance. The first ATC message was presented 3 min later, after participants had lifted off the runway and climbed to 1,200 ft (365.76 m). During the flight, pilots heard 16 ATC messages, presented at the rate of one message every 3 min, directing the pilot to fly a new heading, a new altitude, dial in a new radio frequency, and in 50% of the legs, dial in a new transponder code. Participants were instructed to read back the ATC messages and then execute them in order, according to FAA standards. To further increase workload, pilots were confronted with randomly presented emergency situations: engine malfunctions (i.e., carburetor icing, drop of engine oil pressure) on 8 of 16 flight legs, and/or suddenly approaching air traffic on 10 of 16 flight legs. Pilots were to report engine malfunctions immediately and to avoid conflicting air traffic by veering quickly, yet safely, in the direction diagonal to the path of the oncoming plane. Pilots flew in moderate turbulence throughout the flight and also encountered a 15-knot crosswind during landing approach and actual landing. Crosswinds were always from the right side of the aircraft. The form of the scenarios was very consistent (ATC instructions were recorded by trained controllers), but the scenarios themselves were computer controlled. Several components were randomly varied (e.g. radio frequencies, headings, and altitudes).

Data Analytic Approach

The a priori hypothesis of the study was that an initial measure of processing speed would predict age-related change in aviator performance. However, given the wide variety of cognitive abilities assessed by the CogScreen battery and the large sample size of this study, we broadened the analyses to examine all seven composite measures in the battery. We used a two-step data analytic approach.

First, we performed an exploratory signal detection analysis including all cognitive predictors on two third of the participants to identify potential predictors of age-related change. Second, we conducted a confirmatory mixed effects model analysis (growth curve analysis) using the remaining one third of the participants to examine whether the predic-

tors identified from the first step in fact predict the change. Specifically, in the first step, we employed a receiver operating characteristic (ROC) curve analysis (Kraemer, 1992). In the second step, we employed a random-effect linear growth model (Caselli et al., 2009; Fitzmaurice, Laird, & Ware, 2004; Singer & Willett, 2003), confirming the ROC findings. In this two-step approach, we benefit from integrating the nonparametric ROC approach, which helps identify potential predictors of change without limiting the pool of candidate variables, and the parametric mixed effects growth curve modeling, which is a flexible and efficient tool for modeling change over time. The latter approach is particularly valuable when most subjects in the sample do not have complete data for the entire age span under study, as in our case. This approach allowed us to determine if, as hypothesized, processing speed predicted age-related longitudinal change in flight performance. It also allowed us to determine if other cognitive measures or their interaction with processing speed augments the ability to predict longitudinal change.

ROC curve analysis. —In essence, the ROC searches all the predictor variables (in this study, initial cognitive test scores) and identifies the predictors with the optimal balance between sensitivity and specificity for identifying those aviators with the outcome of interest. The first step is to define the outcome and to choose success/failure criteria. In the current study, we defined the outcome as a decline in the flight summary score that was faster than the median rate of decline. We used a linear mixed model to calculate each individual's rate-of-decline score. Because the purpose of this analysis was to obtain a decline score that would be used as an outcome in the ROC analysis, we did not include any predictor variables in the mixed model. Using the estimated decline scores obtained from a linear mixed model, we conducted an ROC analysis to select the predictor variables of interest. Although many characteristics of aviators might be associated with a more rapid change in function, we focused on the seven CogScreen cognitive composite scores (collected at the baseline visit) because this battery represents measures considered relevant to aviation and are routinely assessed by FAA aeromedical examiners when evaluating pilots about whom there is concern regarding cognitive impairment.
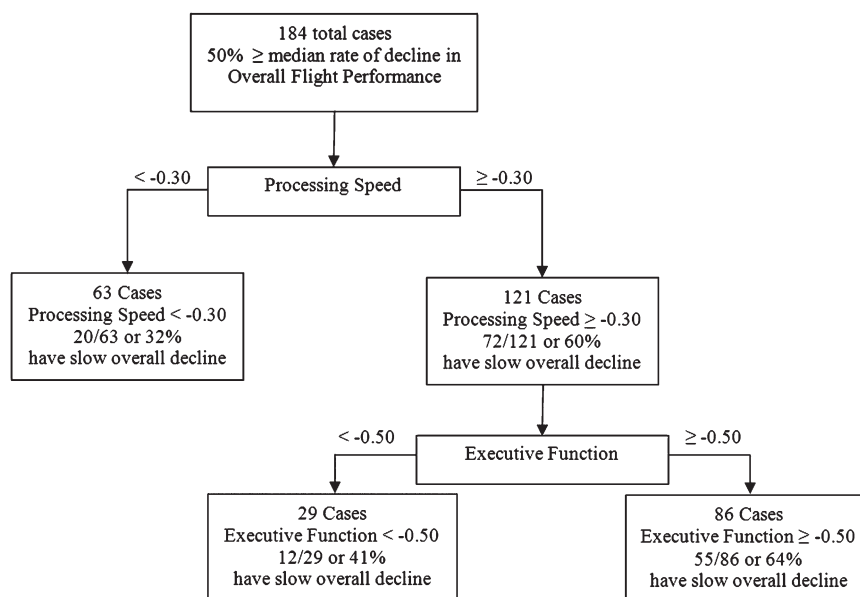
Figure 1. Exploratory signal detection analysis or receiver operator characteristic (ROC) decision tree for predictors of performance greater than or equal to the median rate of decline of summary score on flight simulation task with 184 or two-thirds of all subjects. The analysis starts with 184 pilots, 50% of whom have a rate of decline in overall flight performance slower than the median and 50% have a faster rate of decline than the median, that is, the 184 pilots are characterized as having (or not) at least the median rate of decline in overall performance. The median split of the rate of decline success criterion and the two-third allocation of all subjects allow the maximum possibility of branching of the decision tree in this exploratory analysis. See Results and Discussion for additional explanation.

We used publicly available software we have developed for the ROC analysis (ROC version 4.22; Yesavage & Kraemer, 2007). The ROC software searches all the predictor variables and their associated cut points (for the cognitive test scores) and selects the scores that best identify those aviators with more rapid decline. Once the optimal variable and associated cut points are identified, the association with the success criterion is tested against a stopping rule ($p <$ .05); and/or when a subgroup has too small a sample size for further analysis ($n < 20$); and/or when there are no further variables selected. If the association passes the rule, the sample is divided into two groups according to performance on the optimal variable. The ROC analysis is then restarted, separately, for each of these two sub-groups. The result is a decision tree (Figure 1). As shown in the figure, the signal detection approach identified two cognitive predictors: processing speed and executive function.

Confirmation of ROC findings with mixed model analyses.—To confirm the effects of cognitive processing speed and executive function on longitudinal performance, we used mixed effects growth curve modeling. We assumed a linear trend of performance over time (i.e., age). We also conducted the analysis allowing for a nonlinear trend (by adding an Age × Age term in the model). However, allowing for a nonlinear trend did not improve the fit of the growth model (results not reported). In the growth analysis assuming a linear trend, age (at each study visit) was centered on 60 (which was near the median age of 59.7 years). Three

potential predictors of the longitudinal trend were included in this model: level of expertise, initial processing speed scores, and initial executive function scores. The covariates were tested in relation to both initial performance and age-related change (estimated by the fitted intercepts and slopes, respectively). Expertise was centered on the middle of the three levels of experience, that is, experience was coded as −1 (VFR), 0 (IFR) and +1 (CFII/ATP). Processing speed and executive function scores were centered at the median score at baseline. Centering of these predictor covariates (Kraemer & Blasey, 2004) was important because we were interested in estimating not only the effects of potential predictors on the trend, but also the trend itself, which involves interpretation of the intercept (initial performance) in this analysis. Estimation of our growth curve model was performed using the PROC MIXED procedure in SAS software, version 9.1.3 (Cary, NC). Finally, to obtain an indication of which, if any, flight components accounted for the effects identified by the ROC analysis as predicting performance on the flight summary score, similar mixed model analyses were performed on each of the four flight components that comprise the summary score.

RESULTS

The exploratory ROC analysis suggested that processing speed was the best predictor of age-related longitudinal change in the summary measure of flight simulator performance ($\kappa = .25$; $\chi^2 = 12.77$, $p < .001$). In addition, processing speed and executive function predicted longitudinal

Table 3. Confirmatory Mixed Effects Growth Curve Analysis of Longitudinal Flight Simulator Performance Using Predictors Identified by Exploratory Signal Detection

	Parameter estimate (SE)	p Value
Initial performance (I)		
Intercept (mean, η_I)	0.071 (0.042)	.095
Processing speed (β_{I1})	0.264 (0.074)	.0006
Executive function (β_{I2})	0.065 (0.049)	.188
Expertise (β_{I3})	0.247 (0.058)	<.0001
Processing speed × Executive function (β_{I4})	−0.053 (0.074)	.476
Change in performance over age (S)		
Intercept (mean, η_S)	−0.019 (0.004)	<.0001
Processing speed (β_{S1})	0.021 (0.008)	.010
Executive function (β_{S2})	0.003 (0.006)	.642
Expertise (β_{S3})	0.005 (0.006)	.351
Processing speed × Executive Function (β_{S4})	0.021 (0.008)	.008

Notes: Parameter estimates are based on one-third of the sample. Only 86 subjects were included in this model because six subjects were missing executive function scores. The results reported in Table 3 are based on a model assuming a linear trend of performance over age. Allowing for a nonlinear age trend did not improve the fit of the growth model. The model for the outcome at a given age was: $Y = I + S$ (age centered) $+ e$. The model for initial performance was: $I = \eta_I + \beta_{I1}$ (processing speed centered) $+ \beta_{I2}$ (executive function centered) $+ \beta_{I3}$ (expertise centered) $+ \beta_{I4}$ (processing speed centered) (executive function centered) $+ \zeta_I$. The model for the longitudinal change (slope) of performance was: $S = \eta_S + \beta_S$ (processing speed centered) $+ \beta_{S2}$ (executive function centered) $+ \beta_{S3}$ (expertise centered) $+ \beta_{S4}$ (processing speed centered) (executive function centered).

flight performance in an interactive manner. As shown in Figure 1, the optimal cut point for predicting an above the median (slower) rate of decline is a processing speed *z*-score of greater than or equal to −0.30 (negative z-scores indicate slower cognitive speed). This cut point correctly identified 60% of the 121 pilots that scored above the cut point as having an above median (slower) rate of decline. In contrast, of the 63 pilots that scored lower than this cut

point for processing speed, only 32% had above (slower) median rates of decline. A second-level branching of those who scored faster on processing speed suggested a further advantage for those with a *z*-score better than −.50 on executive function. In other words, slower rate of decline in flight simulator performance was predicted by faster processing speed and better executive function scores ($\kappa = .19$; $\chi^2 = 4.55$, $p < .05$). The other composite CogScreen measures (assessments of episodic memory, working memory, and psychomotor performance) did not predict rate of decline.

In the confirmatory mixed model of longitudinal flight performance, the effects of processing speed and executive function were specifically examined because they were identified by the ROC analysis as potential predictors of change in performance. Results are summarized in Table 3. Note that the outcome measure used in this analysis is the Flight Performance Summary Score. The upper section of Table 3 shows the model's estimates for initial level of performance and predictors of initial performance. It shows that processing speed and expertise significantly influenced initial level of performance. The bottom section of Table 3 shows the relationship between age-related decline in simulator performance and predictors of rate of decline. On average (intercept: η_S), there was a significant decline in simulator performance with increasing age (−.019 *z*-score units per year, $p < .0001$). Importantly, processing speed and the interaction between processing speed and executive function were significant predictors of age-related decline in flight performance. These results can be interpreted as: pilots with faster processing speed and higher executive function scores at baseline had slower rates of decline in flight simulator performance than their same-aged counterparts. Figure 2 graphically illustrates the association between
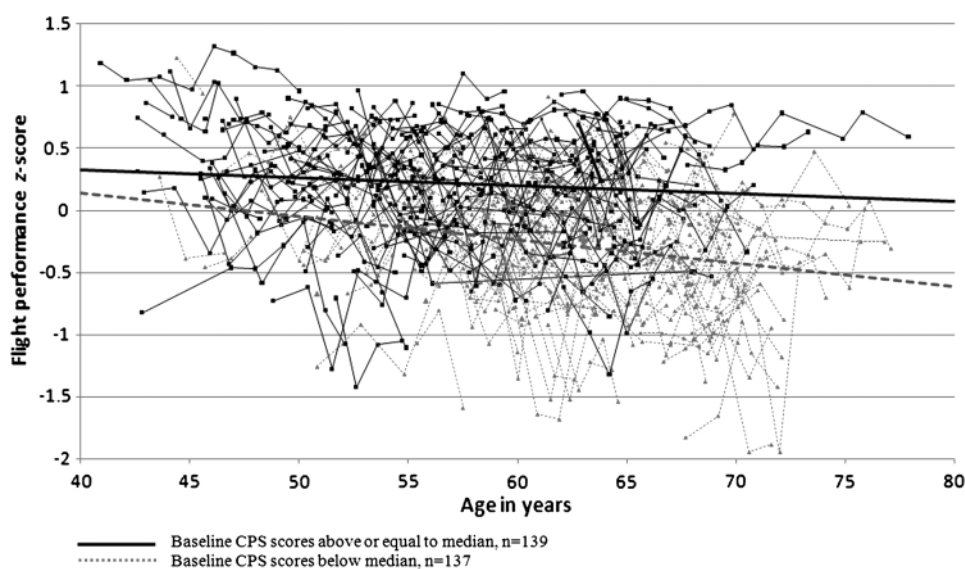

Figure 2. Estimated mean trajectories and observed individual trajectories of flight performance (*z*-score) for aviators with cognitive processing speed (CPS) scores above or below the median baseline for all 276 aviators. Baseline CPS scores above or equal to median, *n* = 139. Baseline CPS scores below median, *n* = 137.

processing speed and longitudinal decline in flight performance. Shown in the figure are the estimated mean trajectories for pilots with above *versus* below the median processing speed scores.

Expertise was not a significant predictor of age-related decline in flight performance (*p* = .35), although the parameter estimate was in the expected direction. Regarding which aspects of change in flight performance components were predicted by covariates in the model, only change in performance on the landing approach component was predicted by initial processing speed (β_{S1} = 0.015; *SE* = 0.007; *p* = .03).

Discussion

Results from the current study suggest that longitudinal flight simulator performance can be predicted by initial assessment of cognitive abilities relevant to flying, that is, processing speed and executive function. These findings are particularly applicable to general aviation such as recreational flying. The percentage of older general aviators has increased in recent years (Nakagawara, Montgomery, & Wood, 2004), and older age (as well as inexperience) has been associated with increased risk of aviation accidents (Li, Baker, Qiang, Grabowski, & McCarthy, 2005). Cross-sectional studies of aviators have demonstrated that measures of basic perceptual and cognitive processing abilities correlate with same-day performance of flight tasks (Kennedy et al., 2010; Morrow et al., 2001; Sohn & Doane, 2004; Taylor et al., 2000); the present study adds to the current literature by demonstrating the feasibility of predicting future flight simulator performance on the basis of initial cognitive measures.

The interaction of processing speed and executive function as predictors of longitudinal performance is of particular interest in studies of a complex task such as aviation performance. In aviation, adequate executive function in terms of "planning ahead" is very important so that the pilot does not "get behind" the aircraft, as critical tasks must be performed correctly in rapid sequence (Wise, Hopkin, & Garland, 2009). Thus, on the types of tasks assessed in the simulator, it would be expected that those pilots who could plan and best allocate cognitive resources would perform well on such tasks.

It may be interesting to contrast the results on the processing speed by executive function interaction in the ROC analyses *versus* the confirmatory mixed effects growth curve analyses. An advantage of ROC analyses is that they may provide more understandable "cutoffs" on scores to interpret results. In the ROC, we note that the relationship shows that the best cutoff for executive function testing is 0.5 *SE* (*z*-scores) below mean performance in the group with relatively fast speed of processing. This implies that the cutoff identifies a subpopulation that despite a relatively high level of performance speed of processing testing performs poorly on executive function testing and that an executive function deficit to this degree predicts poor future performance. In other words, fast speed of processing does not make up for an executive function deficit. Interpreting a parameter estimate for the same interaction from a mixed-effects growth curve analysis may be more challenging.

These results have implications for other cognitively demanding real-world tasks such as driving. It is known that slower processing speed predicts a greater likelihood of driving cessation within 5 years (Anstey, Windsor, Luszcz, & Andrews, 2006; Edwards et al., 2008). Processing speed measures such as the Digit Symbol Substitution test (Wechsler, 1981) predicted driving cessation, both independently and after adjusting for effects of age and self-rated health/physical function (Anstey, Windsor, et al., 2006; Edwards et al., 2008). Additionally, decline of processing speed (as well as general cognitive, verbal, and memory abilities) has been linked to increases in rates of falling and fall risk in elderly participants (Anstey, von Sanden, & Luszcz, 2006; Welmerink, Longstreth, Lyles, & Fitzpatrick, 2010). Because mobility is such a key part of maintaining independence in later life, it would be ideal if brief measures of cognitive function, including processing speed and executive function, were employed to identify individuals at risk and to target the most effective and appropriate intervention. Nonetheless, it is unlikely that purely cognitive tests could fully replace field testing in determining either driving or aviation competency.

The results of this study relate to an expanding literature that emphasizes expertise (Krampe & Ericsson, 1996), cognitive reserve (Stern, 2006), and enrichment (Hertzog, Kramer, Wilson, & Lindenberger, 2009) as factors that moderate (reduce) the impact of aging on cognitive performance. It might seem surprising that higher level of expertise did not significantly slow the rate of age-related change in aviation performance in the current study. However, based on an extensive literature review, Salthouse (2006) also found no Expertise × Age interactions in performance of occupationally relevant cognitive tasks involving individuals in cognitively complex occupations, such as physicians, architects, and college professors. Meinz (2000) noted that the largest group of studies not finding evidence for an age-moderating effect of expertise were studies involving recent episodic memory, for example, recall or reproduction of meaningfully arranged chess pieces (Charness, 1981a, 1981b), musical memory by written recall and by piano keyboard execution (Meinz, 2000; Meinz & Salthouse, 1998), and immediate memory of simulated baseball game broadcasts (Hambrick & Engle, 2002). In summary, it appears that expertise seems at most to preserve only some realms of competence against the effects of age.

That said, it should be emphasized that enriching experiences have significant benefits for "level" of performance, as reported in the present study and other published research. Reaching a higher level of performance in midlife

has benefits in that performance is more likely to maintain function above a given threshold of competence longer. In a similar manner, our results indicate expertise was associated with higher baseline performance of aviation tasks; hence, more expert pilots would maintain a higher level of function longer than less expert pilots, even though both groups' performances declined at the same rate. As Wilson (2009) pointed out, the robust association of education with level of cognitive function primarily accounts for influence of education on lowering risk of dementia in old age. Cognitive training (Hertzog et al., 2009) and continued maintenance of skills through deliberate "maintenance" practice (Krampe & Ericsson, 1996) in later life has been associated with better function in late life. Future longitudinal research should characterize the extent to which an intervention improves level of function and the extent to which the intervention alters the rate of age-related change in function.

Limitations of this study include the relatively brief span of follow-up achieved on average (3.8 years) and suboptimal measures of some constructs. Specifically, the measures of memory and psychomotor ability were based on two-item scores. Also, the episodic memory measure is prone to ceiling effects because this test has only six paired associates (Taylor et al., in press). By comparison, the measures of processing speed and executive function were composites of at least three items. The three-level ordinal measure of expertise based on FAA pilot proficiency ratings is a relatively crude measure. It is possible, for example, that certain pilots better manage their cognitive resources with a disciplined scan of instruments and the external environment. Thus, eye-tracking measures of scanning behavior (Kasarskis et al., 2001; Sarter, Mumaw, & Wickens, 2007; Schriver et al., 2008) may provide a quantitative domain-specific measure of expertise that better predicts flight performance.

In conclusion, the current study suggests that prediction of longitudinal performance of an important real-world activity can be improved by inclusion of initial performance on a relevant cognitive task. This is not to say that age or performance-based criteria should be replaced by a psychometric test. It may suggest that prediction of future performance as a pilot could be bolstered not only by evidence of current adequate performance, but also by psychometric data determining the speed of basic cognitive processes and executive function.

REFERENCES

H.R. 4343–110th Congress: Fair Treatment for Experienced Pilots Act. (2007). *GovTrack.us (database of federal legislation)*. Retrieved July 6, 2010, from http://www.govtrack.us/congress/bill.xpd?bill=h110-4343

Anstey, K. J., von Sanden, C., & Luszcz, M. A. (2006). An 8-year prospective study of the relationship between cognitive performance and falling in very old adults. *Journal of the American Geriatrics Society*, *54*, 1169–1176. doi:10.1111/j.1532-5415.2006.00813.x

Anstey, K. J., Windsor, T. D., Luszcz, M. A., & Andrews, G. R. (2006). Predicting driving cessation over 5 years in older adults: Psychological well-being and cognitive competence are stronger predictors than physical health. *Journal of the American Geriatrics Society*, *54*, 121–126. doi:10.1111/j.1532-5415.2005.00471.x

Bellenkes, A. H., Wickens, C. D., & Kramer, A. F. (1997). Visual scanning and pilot expertise: The role of attentional flexibility and mental model development. *Aviation Space and Environmental Medicine*, *68*, 569–579.

Birren, J. E., & Fisher, L. M. (1995). Aging and speed of behavior: Possible consequences for psychological functioning. *Annual Review of Psychology*, *46*, 329–353. doi:10.1146/annurev.ps.46.020195.001553

Buckner, R. L. (2004). Memory and executive function in aging and AD: Multiple factors that cause decline and reserve factors that compensate. *Neuron*, *44*, 195–208. doi:10.1016/j.neuron.2004.09.006

Bucur, B., Madden, D. J., Spaniol, J., Provenzale, J. M., Cabeza, R., White, L. E., & Huettel, S. A. (2008). Age-related slowing of memory retrieval: Contributions of perceptual speed and cerebral white matter integrity. *Neurobiology of Aging*, *29*, 1070–1079. doi:10.1016/j.neurobiolaging.2007.02.008

Bugg, J. M., Zook, N. A., DeLosh, E. L., Davalos, D. B., & Davis, H. P. (2006). Age differences in fluid intelligence: Contributions of general slowing and frontal decline. *Brain and Cognition*, *62*, 9–16. doi:10.1016/j.bandc.2006.02.006

Caselli, R. J., Dueck, A. C., Osborne, D., Sabbagh, M. N., Connor, D. J., Ahern, G. L., . . . Reiman, E. M. (2009). Longitudinal modeling of age-related memory decline and the APOE epsilon4 effect. *New England Journal of Medicine*, *361*, 255–263. doi:10.1056/NEJMoa0809437

Charness, N. (1981a). Aging and skilled problem solving. *Journal of Experimental Psychology: General*, *110*, 21–38. doi:10.1037/0096-3445.110.1.21

Charness, N. (1981b). Visual short-term memory and aging in chess players. *Journal of Gerontology*, *36*, 615–619. doi:10.1093/geronj/36.5.615

Dempster, F. N. (1992). The rise and fall of the inhibitory mechanism: Toward a unified theory of cognitive development and aging. *Developmental Review*, *12*, 45–75. doi:10.1016/0273-2297(92)90003-K

Doane, S. M., Sohn, Y. W., & Jodlowski, M. T. (2004). Pilot ability to anticipate the consequences of flight actions as a function of expertise. *Human Factors*, *46*, 92–103. doi:10.1518/hfes.46.1.92.30386

Edwards, J. D., Ross, L. A., Ackerman, M. L., Small, B. J., Ball, K. K., Bradley, S., & Dodson, J. E. (2008). Longitudinal predictors of driving cessation among older adults from the ACTIVE clinical trial.

The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, *63*, 6–12.

Finkel, D., Reynolds, C. A., McArdle, J. J., Hamagami, F., & Pedersen, N. L. (2009). Genetic variance in processing speed drives variation in aging of spatial and memory abilities. *Developmental Psychology*, *45*, 820–834. doi:10.1037/a0015332

Fitzmaurice, G. M., Laird, N. M., & Ware, J. H. (2004). *Applied longitudinal analysis*. New York: John Wiley and Sons.

Friedman, N. P., Miyake, A., Young, S. E., Defries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*, 201–225. doi:10.1037/0096-3445.137.2.201

Gray, J. R., Chabris, C. F., & Braver, T. S. (2003). Neural mechanisms of general fluid intelligence. *Nature Neuroscience*, *6*, 316–322. doi:10.1038/nn1014

Gunning-Dixon, F. M., & Raz, N. (2003). Neuroanatomical correlates of selected executive functions in middle-aged and older adults: A prospective MRI study. *Neuropsychologia*, *41*, 1929–1941. doi:10.1016/S0028-3932(03)00129-5

Hambrick, D. Z., & Engle, R. W. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology*, *44*, 339–387. doi:10.1006/cogp.2001.0769

Hasher, L. & Zacks, R. T. (Eds.), (1988). *Working memory, comprehension, and aging: A review and a new view* (Vol. 22). New York: Academic Press. doi:10.1016/S0079-7421(08)60041-9

Hasher, L., Zacks, R. T., & May, C. P. (1999). Inhibitory control, circadian arousal, and age. In D. Gopher & A. Koriat (Eds.), *Attention and performance XVII: Cognitive regulation of performance: Interaction of theory and application* (pp. 653–675). Cambridge, MA: MIT Press.

Head, D., Kennedy, K. M., Rodrigue, K. M., & Raz, N. (2009). Age differences in perseveration: Cognitive and neuroanatomical mediators of performance on the Wisconsin Card Sorting Test. *Neuropsychologia*, *47*, 1200–1203. doi:10.1016/j.neuropsychologia.2009.01.003

Hedden, T., & Park, D. (2001). Aging and interference in verbal working memory. *Psychology and Aging*, *16*, 666–681. doi:10.1037/0882-7974.16.4.666

Hertzog, C., Dixon, R. A., Hultsch, D. F., & MacDonald, S. W. (2003). Latent change models of adult cognition: Are changes in processing speed and working memory associated with changes in episodic memory? *Psychology and Aging*, *18*, 755–769.

Hertzog, C., Kramer, A. F., Wilson, R. S., & Lindenberger, U. (2009). Enrichment effects on adult cognitive development: Can the functional capacity of older adults be preserved and enhanced? *Psychological Science in the Public Interest*, *9*, 1–65.

Kasarskis, P., Stehwien, J., Hickox, J., Aretz, A., & Wickens, C. D. (2001). *Comparison of expert and novice scan behaviors during VFR fight* Paper presented at the 11th International Symposium on Aviation Psychology. Columbus: The Ohio State University.

Kay, G. G. (1995). *CogScreen Aeromedical edition professional manual*. Odessa, FL: Psychological Assessment Resources.

Kennedy, Q., Taylor, J. L., Reade, G., & Yesavage, J. A. (2010). Age and expertise effects in aviation decision making and flight control in a flight simulator. *Aviation Space and Environmental Medicine*, *81*, 489–497. doi:10.3357/ASEM.2684.2010

Kraemer, H. C. (1992). *Evaluating medical tests: Objective and quantitative guidelines*. Newbury Park, CA: Sage.

Kraemer, H. C., & Blasey, C. M. (2004). Centring in regression analyses: A strategy to prevent errors in statistical inference. *International Journal of Methods in Psychiatric Research*, *13*, 141–151. doi:10.1002/mpr.170

Krampe, R. T., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General*, *125*, 331–359. doi:10.1037/0096-3445.125.4.331

Kwong See, S. T., & Ryan, E. B. (1995). Cognitive mediation of adult age differences in language performance. *Psychology and Aging*, *10*, 458–468. doi:10.1037/0882-7974.10.3.458

Lassiter, D. L., Morrow, D. G., Hinson, G. E., Miller, M., & Hambrick, D. Z. (1996). Expertise and age effects on pilot mental workload in a simulated aviation task. *Human Factors and Ergonomics Society Annual Meeting Proceedings*, *40*, 133–137.

Li, G., Baker, S. P., Qiang, Y., Grabowski, J. G., & McCarthy, M. L. (2005). Driving-while-intoxicated history as a risk marker for general aviation pilots. *Accident Analysis and Prevention*, *37*, 179–184. doi:10.1016/j.aap.2004.04.005

Lindenberger, U., Mayr, U., & Kliegl, R. (1993). Speed and intelligence in old age. *Psychology and Aging*, *8*, 207–220. doi:10.1037/0882-7974.8.2.207

Meinz, E. J. (2000). Experience-based attenuation of age-related differences in music cognition tasks. *Psychology and Aging*, *15*, 297–312. doi:10.1037/0882-7974.15.2.297

Meinz, E. J., & Salthouse, T. A. (1998). The effects of age and experience on memory for visually presented music. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *53*, 60–69. doi:10.1093/geronb/53B.1.P60

Morrow, D. G., Leirer, V., Altieri, P., & Fitzsimmons, C. (1994). When expertise reduces age differences in performance. *Psychology and Aging*, *9*, 134–148. doi:10.1037/0882-7974.9.1.134

Morrow, D. G., Menard, W. E., Stine-Morrow, E. A., Teller, T., & Bryant, D. (2001). The influence of expertise and task factors on age differences in pilot communication. *Psychology and Aging*, *16*, 31–46. doi:10.1037/0882-7974.16.1.31

Morrow, D. G., Miller, L. M., Ridolfo, H. E., Menard, W., Stine-Morrow, E. A., & Magnor, C. (2005). Environmental support for older and younger pilots' comprehension of air traffic control information. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, *60*, 11–18. doi:10.1093/geronb/60.1.P11

Morrow, D. G., Ridolfo, H. E., Menard, W. E., Sanborn, A., Stine-Morrow, E. A., Magnor, C., . . . Bryant, D. (2003). Environmental support promotes expertise-based mitigation of age differences on pilot communication tasks. *Psychology and Aging*, *18*, 268–284. doi:10.1037/0882-7974.18.2.268

Nakagawara, V. B., Montgomery, R. W., & Wood, K. J. (2004). *Demographics and vision restrictions in civilian pilots: Clinical implications*. Oklahoma City, OK: Civil Aerospace Medical Institute.

Salthouse, T. A. (1991). Mediation of adult age differences in cognition by reductions in working memory and speed of processing. *Psychological Science*, *2*, 179–183. doi:10.1111/j.1467-9280.1991.tb00127.x

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review*, *103*, 403–428. doi:10.1037/0033-295X.103.3.403

Salthouse, T. A. (2006). Mental exercise and mental aging: Evaluating the validity of the "Use it or Lose it" hypothesis. *Perspectives on Psychological Science*, *1*, 68–87. doi:10.1111/j.1745-6916.2006.00005.x

Sarter, N. B., Mumaw, R. J., & Wickens, C. D. (2007). Pilots' monitoring strategies and performance on automated flight decks: An empirical study combining behavioral and eye-tracking data. *Human Factors*, *49*, 347–357. doi:10.1518/001872007X196685

Schaie, K. W. (1989). Perceptual speed in adulthood: Cross-sectional and longitudinal studies. *Psychology and Aging*, *4*, 443–453. doi:10.1037/0882-7974.4.4.443

Schretlen, D., Pearlson, G. D., Anthony, J. C., Aylward, E. H., Augustine, A. M., Davis, A., & Barta, P. (2000). Elucidating the contributions of processing speed, executive ability, and frontal lobe volume to normal age-related differences in fluid intelligence. *Journal of the International Neuropsychological Society*, *6*, 52–61. doi:10.1017/S1355617700611062

Schriver, A. T., Morrow, D. G., Wickens, C. D., & Talleur, D. A. (2008). Expertise differences in attentional strategies related to pilot decision making. *Human Factors*, *50*, 864–878. doi:10.1518/001872008X374974

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence*. New York: Oxford University Press. doi:10.1093/acprof:oso/9780195152968.003.0002

Sohn, Y. W., & Doane, S. M. (2004). Memory processes of flight situation awareness: Interactive roles of working memory capacity, long-term working memory, and expertise. *Human Factors*, *46*, 461–475. doi:10.1518/hfes.46.3.461.3782

Stern, Y. (2006). Cognitive reserve and Alzheimer disease. *Alzheimer Disease and Associated Disorders*, *20*(3 Suppl. 2),S69–S74. doi:10.1097/00002093-200607001-00010

Taylor, J. L., Kennedy, Q., Adamson, M. M., Lazzaroni, L., Noda, A., Murphy, G. M., & Yesavage, J. A. Influences of APOE ε4 and expertise on performance of older pilots. *Psychology and Aging*. Advanced online publication. doi:10.1037/a0021697

Taylor, J. L., Kennedy, Q., Noda, A., & Yesavage, J. A. (2007). Pilot age and expertise predict flight simulator performance: A 3-year longitudinal study. *Neurology*, *68*, 648–654. doi:10.1212/01.wnl. 0000255943.10045.c0

Taylor, J. L., O'Hara, R., Mumenthaler, M. S., Rosen, A. C., & Yesavage, J. A. (2005). Cognitive ability, expertise, and age differences in following air-traffic control instructions. *Psychology and Aging*, *20*, 117–133. doi:10.1037/0882-7974.20.1.117

Taylor, J. L., O'Hara, R., Mumenthaler, M. S., & Yesavage, J. A. (2000). Relationship of CogScreen-AE to flight simulator performance and pilot age. *Aviation, Space, and Environmental Medicine*, *71*, 373–380.

Tsang, P. S., & Shaner, T. L. (1998). Age, attention, expertise, and time-sharing performance. *Psychology and Aging*, *13*, 323–347. doi:10.1037/0882-7974.13.2.323

Walhovd, K. B., Fjell, A. M., Reinvang, I., Lundervold, A., Fischl, B., Salat, D., . . . Dale, A. M. (2005). Cortical volume and speed-of-processing are complementary in prediction of performance intelligence. *Neuropsychologia*, *43*, 704–713. doi:10.1016/j.neuropsychologia. 2004.08.006

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale—revised manual*. New York: Psychological Corporation.

Welmerink, D. B., Longstreth, W. T., Jr., Lyles, M. F., & Fitzpatrick, A. L. (2010). Cognition and the risk of hospitalization for serious falls in the elderly: Results from the cardiovascular health study. *The Journals of Gerontology, Series A: Biological Sciences and Medical Sciences*, *65*, 1242–1249. doi:10.1093/gerona/glq115

West, R. L. (1996). An application of prefrontal cortex function theory to cognitive aging. *Psychological Bulletin*, *120*, 272–292. doi:10.1037/0033-2909.120.2.272

Wiggins, M., & O'Hare, D. (1995). Expertise in aeronautical weather-related decision making: A cross-sectional analysis of general aviation pilots. *Journal of Experimental Psychology: Applied*, *1*, 305–320. doi:10.1037/1076-898X.1.4.305

Wilson, R. S., Beckett, L. A., Barnes, L. L., Schneider, J. A., Bach, J., Evans, D. A., & Bennett, D. A. (2002). Individual differences in rates of change in cognitive abilities of older persons. *Psychology and Aging*, *17*, 179–193. doi:10.1037/0882-7974.17.2.179

Wilson, R. S., Hebert, L. E., Scherr, P. A., Barnes, L. L., Mendes de Leon, C. F., & Evans, D. A. (2009). Educational attainment and cognitive decline in old age. Neurology, *72*(5), 460–465. doi:10.1212/01. wnl.0000341782.71418.6c

Wise, J. A., Hopkin, V. D., & Garland, D. J. (2009). *Handbook of aviation human factors*. Boca Raton, FL: CRC Press. doi:10.1201/b10401-28

Yesavage, J. A., & Kraemer, H. C. (2007). *Signal Detection Software for Receiver Operator Characteristics (Version 4.22)*. [Computer software]. Palo Alto, CA. Retrieved from http://www.stanford.edu/~ yesavage/ROC.html

Yesavage, J. A., Taylor, J. L., Mumenthaler, M. S., Noda, A., & O'Hara, R. (1999). Relationship of age and simulated flight performance. *Journal of the American Geriatrics Society*, *47*, 819–823.

Zimprich, D., & Martin, M. (2002). Can longitudinal changes in processing speed explain longitudinal age changes in fluid intelligence? *Psychology and Aging*, *17*, 690–695. doi:10.1037/0882-7974.17.4.690