



Published in final edited form as:

*Neurology*. 2007 February 27; 68(9): 648–654. doi:10.1212/01.wnl.0000255943.10045.c0.

## Pilot age and expertise predict flight simulator performance:

### A 3-year longitudinal study

Joy L. Taylor, PhD, Quinn Kennedy, PhD, Art Noda, MS, and Jerome A. Yesavage, MD

From the Department of Veterans Affairs and Sierra-Pacific MIRECC (J.L.T., J.A.Y.), Palo Alto; and Stanford University School of Medicine (J.L.T., Q.K., A.N., J.A.Y.), Stanford, CA.

### Abstract

**Background**—Expert knowledge may compensate for age-related declines in basic cognitive and sensory-motor abilities in some skill domains. We investigated the influence of age and aviation expertise (indexed by Federal Aviation Administration pilot ratings) on longitudinal flight simulator performance.

**Methods**—Over a 3-year period, 118 general aviation pilots aged 40 to 69 years were tested annually, in which their flight performance was scored in terms of 1) executing air-traffic controller communications; 2) traffic avoidance; 3) scanning cockpit instruments; 4) executing an approach to landing; and 5) a flight summary score.

**Results**—More expert pilots had better flight summary scores at baseline and showed less decline over time. Secondary analyses revealed that expertise effects were most evident in the accuracy of executing aviation communications, the measure on which performance declined most sharply over time. Regarding age, even though older pilots initially performed worse than younger pilots, over time older pilots showed less decline in flight summary scores than younger pilots. Secondary analyses revealed that the oldest pilots did well over time because their traffic avoidance performance improved more vs younger pilots.

**Conclusions**—These longitudinal findings support previous cross-sectional studies in aviation as well as non-aviation domains, which demonstrated the advantageous effect of prior experience and specialized expertise on older adults' skilled cognitive performances.

As the workforce ages in an era of accelerating technological advances, it becomes imperative to understand how aging affects performance in the workplace. In aviation, for example, an aging workforce coincident with the introduction of jet aircraft appears to have played a role in the Federal Aviation Administration's (FAA) decision for mandatory retirement of airline pilots at age 60.<sup>1</sup> It has been argued that age-based retirement rules are discriminatory and should be replaced with more direct methods of risk assessment.<sup>2</sup> Simulations of occupationally relevant or hazardous activities such as driving are desirable complements to medical and neuropsychological assessments because simulations permit individuals to draw upon prior knowledge and procedural memory relevant to a skill domain. Cross-sectional studies of expert performers, including medical technologists, typists, and musicians, have found that expert knowledge may compensate for age-related

---

Copyright © 2007 by AAN Enterprises, Inc.

Correspondence to: Joy L. Taylor.

Address correspondence and reprint requests to Dr. Joy L. Taylor, Aging Clinical Research Center, 3801 Miranda Avenue (151Y), Palo Alto, CA 94304; joyt@stanford.edu.

Disclosure: The authors report no conflicts of interest.

**Reprints** Information about ordering reprints can be found online: <http://www.neurology.org/misc/reprints.shtml>

declines in basic cognitive and sensory-motor abilities in some skill domains.<sup>3–8</sup> Thus, expertise may moderate (reduce) the impact of age on occupationally relevant performance.

Flight simulator assessments provide objective, reliable performance measures that are sensitive to differences in age<sup>9</sup> and level of aviation expertise,<sup>10</sup> but all of the flight simulator and expertise research to date has been cross-sectional. Longitudinal studies are essential toward understanding the aging process and its interplay with putative protective factors such as expertise.<sup>11,12</sup> Using data from the ongoing Stanford/VA longitudinal study of aviators aged 40 to 69 years at study entry, we investigated the influences of age and expertise on flight simulator performance over a 3-year period.

## Methods

### Participants

**Entry criteria**—Participants were part of the ongoing longitudinal Stanford/VA Aviation Study approved by the Stanford University Institutional Review Board. Main inclusion criteria were age 40 to 69 years at study entry, current FAA medical certificate (Class III or higher), and current flying activity with 300 to 15,000 hours of total flight time. This range of total flight hours was designed to avoid strong collinearity between age and hours of aviation experience; older airline pilots, for example, typically have over 20,000 hours of total flight time. Retired pilots from major air carriers were excluded because decline in flight simulator performance could be explained by less opportunity to fly after retirement. Thus, we selected a group of pilots whose aviation activity did not necessarily change at age 60. All participants gave written informed consent to participate in annual testing, with the right to withdraw at any time.

At entry, each participant was classified into one of three levels of aviation expertise depending on which FAA pilot proficiency ratings had been previously attained: 1) least expertise: VFR (rated for flying under visual flight rules only); 2) moderate expertise: IFR (also rated for instrument flight); and 3) most expertise: CFII, ATP, or both CFII and ATP (certified flight instructor of IFR students or rated for flying air-transport planes). FAA ratings are a convenient yet valid indicator of expertise level because each rating requires progressively more advanced training and more hours of flight experience. Within the VFR group, all were recreational pilots, though two had aviation-related employment (airplane broker and aircraft mechanic). Within the IFR group, the majority (55/60) had careers unrelated to aviation, though a few were part-time CFIs,<sup>2</sup> aviation analysts,<sup>2</sup> or had been an aviator in the army. Approximately one half (14/26) of the CFII/ATP participants were employed as full-time air transport pilots,<sup>3</sup> part-time air transport pilots,<sup>4</sup> CFIs,<sup>3</sup> or their job duties included aircraft piloting.<sup>4</sup>

Participants completed a cognitive battery designed to test abilities relevant for piloting aircraft, including tests from the CogScreen-AE battery<sup>13</sup> and tests of information processing speed<sup>14</sup> (see table 1 for means and table E-1 on the *Neurology* Web site at [www.neurology.org](http://www.neurology.org) for descriptions of the measures). Participants with at least three annual time points of flight simulator testing were included in the longitudinal data analyses. Of 141 participants who completed baseline testing before June 1, 2001, 118 had at least three annual time points (mean = 3.8, SD = 0.43), representing an average span of 3.1 years of follow-up (SD = 0.6). Of the 23 participants who had fewer than three annual time points, 12 discontinued participation after the baseline visit (8% of 141); 10 discontinued after the first follow-up (7%); and 1 had only two time points due to missing the first annual follow-up (1%). Stepwise logistic regression modeling did not identify any participant characteristics indicative of selective attrition. The characteristics included in the model were age at entry, expertise group membership, years of education, total hours of flight time, gender, necessity

of a FAA medical waiver, self-reported health, performance on five cognitive tests, and overall performance in the flight simulator at entry.

**Demographic and cognitive ability characteristics at entry**—Table 1 summarizes characteristics of the longitudinal participants at entry, separated by expertise group. As shown in the table, the groups differed in mean age [ $F(2,117) = 5.03; p < 0.01$ ]. Also, higher levels of expertise were associated with more total flight time [ $p < 0.0001$ ; nonparametric Kruskal-Wallis  $F(2,117) = 35.30$ ] as would be expected, and with more recent flight time [ $p < 0.01$ ; Kruskal-Wallis  $F(2,117) = 5.30$ ]. We detected no differences in cognitive test scores by expertise group ( $ps > 0.05$ ; effect sizes [ES] ranged from  $-0.18$  to  $0.02$ ). Older age was associated with lower cognitive test scores (all  $ps < 0.01$ , ESs ranged from  $-0.26$  to  $-0.50$ ), which is consistent with previous findings for the early enrollees of this study<sup>10,15</sup> (supplementary data E-1 on the *Neurology* Web site at [www.neurology.org](http://www.neurology.org) lists results of models testing the effects of age, expertise, and their interaction on cognitive test scores). Finally, despite capping total flight time to avoid collinearity between age and flight time, there was a small correlation between age and total flight time ( $r_s = 0.29, p < 0.01$ ).

## Equipment

Pilots “flew” in a Frasca 141 flight simulator (Urbana, IL). The simulator was linked to a computer specialized for graphics (Silicon Graphics, Mountain View, CA) that generated a “through-the-window” visual environment and continuously collected data concerning the aircraft’s position and communication frequencies. This system simulated flying a small single-engine aircraft with fixed landing gear and fixed propeller above flat terrain with surrounding mountains and clear skies. A cockpit speaker system was used to present prerecorded audio messages that simulated an air-traffic controller speaking to the pilot.

## Procedures

Prior to longitudinal data collection, participants had six practice flights in the simulator to gain familiarity with the flight scenario used throughout the study. Participants typically completed their practice 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-minute flight in the morning and a 75-minute flight in the afternoon. Each flight was followed by a 40- to 60-minute battery of cognitive tests. The entire test day lasted approximately 6 hours, including a 40- to 60-minute lunch break. Each flight began with the air-traffic controller’s takeoff clearance. The first air-traffic control (ATC) message was presented 3 minutes 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 minutes, 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 and according to FAA standards. To further increase workload, pilots were confronted with randomly presented emergency situations: engine malfunctions (carburetor icing, drop of engine oil pressure in 8/16 legs), and suddenly approaching air traffic (10/16 legs). Pilots were to report engine malfunctions immediately and to avoid air traffic by veering quickly yet safely in the direction diagonal to the path of the oncoming plane. Pilots flew in severe turbulence throughout the flight, and also encountered a 15-knot crosswind during approach and landing. Multiple versions of this flight scenario were presented to reduce learning of specific maneuvers and ATC items.

## Measures

The scoring system of the flight simulator-computer system produces 23 variables<sup>9,16</sup> 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 baseline visit mean and SD (scores on the morning and afternoon flights were averaged).

The z-scores on the individual measures were aggregated on the basis of previous principal component analyses into four component measures: 1) accuracy of executing the ATC communications; 2) traffic avoidance; 3) scanning cockpit instruments to detect engine emergencies; and 4) executing a visual approach to landing.<sup>9,16</sup>

## Statistical analysis

The four component measures were averaged to create a flight summary score, which was the primary measure of performance. To elucidate further how performance changed over time, the four component measures were analyzed as secondary measures. Random effects modeling was used to examine baseline levels and annual rates of change in the primary and secondary flight measures. For each participant and measure, the participant's scores from each test day were regressed on the age at test, yielding intercept and slope values for the primary summary measure and for each of the four secondary measures. Thus, each participant had five baseline scores (the intercepts at entry age) and five rate-of-change scores (the slopes). To test hypotheses regarding age and expertise, the baseline and rate-of-change scores were analyzed using general linear modeling (SAS Proc GLM). A separate GLM was constructed for the primary and each secondary outcome measure. The terms of these GLMs were intercept, expertise, age, and the age  $\times$  expertise interaction. Expertise was coded as an ordinal variable (-1, 0, 1) and age was a continuous variable centered at the median.<sup>17</sup> The hypotheses regarding the primary outcomes were as follows:

1. The three levels of aviation expertise would be ordinally related to flight summary scores, such that the most expert pilots (CFII/ATP) would have the best baseline scores and the least decline over time.
2. Increased age would be associated with lower baseline scores and steeper decline over time.
3. Higher levels of expertise would benefit older pilots' baseline and longitudinal flight summary scores to a greater extent than younger pilots' scores.

## Results

Table 2 lists the GLM estimates for the model terms (i.e., intercept, age, expertise, and age  $\times$  expertise terms). Flight summary scores modestly declined an average of 0.025 standard units per year ( $p < 0.05$ ; ES = -0.22). The average rates of change in the various component measures of simulator performance varied widely (see intercept term  $\beta_0$  estimates listed in table 2). Communication task performance declined the most steeply, showing a decline of 0.091 units per year ( $p < 0.0001$ ; ES = -0.61). The average rate of decline in visual approach-to-landing performance was modest ( $p < 0.01$ ; ES = -0.26). In contrast, traffic avoidance performance showed an improvement of 0.041 units per year ( $p < 0.05$ ; ES = 0.19). There was virtually no change in the average time to report engine emergencies ( $t < 1$ ; ES = 0.08).

## Effects of expertise and age on flight simulator performance

Beneficial effects of expertise were observed at baseline and longitudinally. The expected age differences at baseline were observed; yet, the longitudinal age patterns were quite different than expected. The GLM estimates are summarized in table 2, and described in detail below.

**Expertise**—Advanced flight ratings and certifications were associated with better flight summary scores at baseline ( $\beta_2 = 0.155, p < 0.05, ES = 0.24$ ) and less decline over time ( $\beta_2 = 0.039, p < 0.05, ES = 0.23$ ). The average performance of CFII/ATP rated pilots was essentially flat over the duration of follow-up (mean slope of the flight summary score =  $0.002 \pm 0.10$ ). VFR-rated pilots had the steepest rate of decline in flight summary scores (mean =  $-0.066 \pm 0.13$ ). IFR-rated pilots had an intermediate rate of change (mean =  $-0.015 \pm 0.11$ ). The beneficial effects associated with aviation expertise were especially apparent in the communication task (see table 2).

Because the three expertise groups differed significantly in terms of hours of flight time, it was important to examine the extent to which more hours of flight experience could also account for better flight simulator performance. Also, the differing amounts of flight time for VFR, IFR, and CFII/ATP pilots progressively widened over the 3 years of follow-up: VFR pilots accumulated an average of  $63.0 \pm 88.4$  hours per year, IFR pilots an average of  $99.6 \pm 94.9$  hours per year, and CFII/ATP pilots an average of  $223.9 \pm 203.4$  hours [ $F(2,117) = 12.69, p < 0.0001$ ]. To examine the role of flight time on pilot performance, we recomputed the baseline and longitudinal age  $\times$  expertise models, replacing FAA pilot ratings with flight time. In the model of baseline performance, the total hours of flight time reported at study entry was tested (along with age and its interaction with flight time). More total flight time did not predict better flight summary scores at baseline ( $p > 0.10; ES = 0.14$ ; GLM parameter estimates are listed in table E-3). Similarly, greater accumulation of hours of flight experience during follow-up was not associated with less longitudinal decline in the flight summary score ( $p > 0.50; ES = 0.04$ ; see table E-3). In short, expertise—defined by advanced training and extensive time engaged in the activity—was a stronger predictor of skilled performance than amount of activity alone. These findings illustrate how expertise is distinct from amount of activity, even though the two may be intercorrelated.

**Age**—Older age was associated with lower flight summary scores at baseline ( $\beta_1 = -0.038, p < 0.0001, ES = -0.58$ ). The effects of age on baseline performance were most evident in the traffic avoidance ( $ES = -0.60$ ) and approach measures ( $ES = -0.47$ ), though age-related differences were significant for all of the flight component measures (see table 2). Longitudinal analysis of the flight summary scores revealed, surprisingly, that older pilots showed less decline over time than younger pilots ( $\beta_1 = 0.004, p < 0.01, ES = 0.25$ ). The unexpected longitudinal age pattern primarily reflects the fact that older pilots improved their traffic avoidance performance more so than younger pilots ( $p < 0.01; ES = 0.25$ ).

To illustrate the age trends, pilots were subgrouped into three age ranges: 40 to 49, 50 to 59, or 60 to 69 years of age at study entry. This grouping reveals that, in terms of overall flight simulator performance, pilots aged 40 to 49 had a mean rate of decline of  $-0.057$  standard units per year; pilots aged 50 to 59 had a mean decline of  $-0.040$  units per year; 60- to 69-year-old pilots had a mean improvement of  $0.018$  units per year. The figure illustrates the age-related and expertise-related patterns of performance over time. Plotted are the baseline means and the directionality of annual change for pilots within the three age ranges and as a function of FAA rating. As can be seen in the figure, the annual rate of decline decreased with increasing age and with increasing expertise. We did not detect an interaction between

age and expertise ( $p = 0.15$ ;  $ES = 0.14$ ). It should be noted that the numbers of participants in the extremes of the age and expertise distributions were modest (5 to 9).

## Discussion

Findings confirm that flight simulator assessments can detect changes in performance related to age and expertise. Over a 3-year span of testing, we observed a significant though modest decline in overall performance, which varied depending on pilot age and FAA proficiency ratings. Of the four flight components assessed, communication task performance declined the most steeply over time. The present study focused on general aviation pilots due to the difficulty in drawing conclusions about age-related performance differences among airline pilots because mandatory retirement impacts the amount and type of flight experience after age 60. Nonetheless, the population of older general aviation pilots is important in its own right because of needs for medical monitoring and because general aviation accident rates have historically been as much as 90 times the rate for air carriers.<sup>18</sup> Remarkably, a recent epidemiologic study reported an increased risk of general aviation accidents with increasing age, beginning at age 35.<sup>18</sup> Hours of flight experience, a variable more accessible than FAA ratings in aviation databases, is consistently found to be a relevant factor in epidemiologic studies of aircraft accident rates.<sup>18,19</sup>

In this study, more expert pilots, i.e., those with advanced FAA pilot ratings and certifications, had better baseline flight simulator performance, especially in the communication and approach-to-landing components. Several cross-sectional studies have documented the advantage of aviation expertise (and hours of flight experience) in laboratory studies of cockpit scanning,<sup>20</sup> processing ATC communications,<sup>8,10,21</sup> performing instrument flight maneuvers,<sup>22</sup> and making weather-related decisions.<sup>23</sup> More expert pilots in this study also showed less decline over time on average. This longitudinal result bolsters previous cross-sectional findings in aviation as well as non-aviation domains, which demonstrated the advantageous effect of prior experience and specialized expertise on older adults' skilled performances.<sup>3-7</sup>

The prevailing theoretical view is that the acquisition of expertise typically requires a decade or more of deliberate, well-structured practice in a particular skill domain (such as music, athletics, or chess)<sup>24</sup> and reflects brain plasticity,<sup>25</sup> such that experts build an elaborate, integrated base of declarative and procedural knowledge. This specialized base of knowledge supports attention to key relationships between individual items of information,<sup>20</sup> anticipation of likely future events,<sup>26</sup> and coordination of motor movements<sup>4</sup> to respond faster and more accurately. For example, expert pilots attend to the relationship between speed and direction of visual information to anticipate the ideal flight path, whereas novices have not attained this skill.<sup>27</sup> Finally, expert knowledge has been characterized as an example of crystallized intelligence, which is more stable across the lifespan<sup>6</sup> than fluid abilities such as episodic memory recollection and executive control.

Aviation expertise was associated with less decline in flight simulator performance over time. Multiple, interrelated mechanisms related to amount and type of aviation knowledge and frequency of use may explain this finding. We conjecture that in addition to drawing upon aviation knowledge, pilots also learned test-taking strategies specific to the flight simulator testing scenario (especially during the practice sessions). To the extent that memory for test-taking strategies fades at annual follow-up visits, pilots with basic ratings might show decline in overall flight simulator performance. In contrast, strategy recollection may be less consequential for pilots with advanced ratings. First, a pilot with advanced ratings has the benefit of a more elaborate base of knowledge. Second, this knowledge base is better adapted<sup>24</sup> to skills measured in the flight simulator, such as precise altitude control.

Third, between the annual tests, a pilot with advanced ratings may engage in flight activities that require continual access of knowledge and practice of skills (e.g., flight instructing or instrument flying to maintain close altitudes and precision runway approaches), which may help maintain some skills measured in the simulator. VFR-rated pilots are less likely to be engaged in such precise flying between annual tests. Future research should record the time spent in specific types of flight activities to address questions of how much and what types of experience promote stable or improved aviation performance.

The age differences in flight simulator performance observed at baseline are consistent with earlier cross-sectional studies, which also found that older pilots executed air-traffic controller communications less accurately on average, evaded air-traffic conflicts less adroitly, and less skillfully approached the runway for landing.<sup>9,10,28</sup> Unexpectedly, older pilots showed less longitudinal decline in overall flight performance than younger pilots. Secondary analyses revealed that the older pilots did well over time in part because their traffic avoidance performance improved more than younger pilots. There are several possible reasons why older pilots maintained their levels of performance over time, including sampling bias related to hardy survivor and nonrandom drop-out effects, birth cohort differences, and truncated age range of the sample. In view of the lack of evidence for nonrandom drop-out biasing, we focus on three other possible explanations: floor effects, in which poor performers have less room to decline than high performers; regression to the mean, in which over time lower performers improve and higher performers decline; and differential practice effects, in which older participants benefit more from repeated testing than the younger participants.

Older pilots performed worse than younger pilots at baseline on average, and therefore, may have less room to decline due to floor effects. A floor effect has been noted for the transponder item of the communication measure.<sup>29</sup> Nevertheless, older and less expert pilots' communication performance continued to decline over time. In other measures, such as approach, it was possible to have very large deviations from ideal positions, and therefore, substantial room to decline. Thus, floor effects do not appear to be a convincing explanation for the finding of less decline on average for older participants.

Regression to the mean may partially explain the findings, particularly in the traffic avoidance task. Two conditions that together allow regression to the mean are unreliable measures and the differential selection in a pre-post design of participants who initially scored at the extremes. Because we did not use an extreme groups pre-post design, regression to the mean is not an obvious explanation. Also, the flight summary score showed excellent consistency over time (intraclass correlation or ICC = 0.79) and reliability was enhanced by having three to four annual points per participant. Importantly, regression to the mean cannot explain, in the case of the summary and the communication scores, why expertise would give rise to higher baseline scores and less decline over time. Nonetheless, the traffic avoidance results may reflect regression to the mean to some extent because traffic avoidance was the component that showed the largest negative age relations at baseline, significant age-related improvement over time, and the least consistency over time (ICC = 0.48).

Another explanation that is consistent with the pattern of results is an age difference in practice effects. Practice effects have been increasingly recognized in longitudinal studies of normal aging and preclinical dementia.<sup>30–33</sup> Indeed, the incremental increase of taking a test the second time has been estimated to be as much as 10 to 15 times larger than the effect of aging 1 year.<sup>30,32</sup> We attempted to minimize practice effects by familiarizing pilots with the simulator scenario prior to the baseline assessment. Nevertheless, performance of the traffic avoidance task continued to improve over time, with older pilots improving more than

younger pilots. There currently are two lines of evidence, in computerized testing, for greater improvement among older adults. Older adults have shown greater improvements than younger adults in reaction time in the task-switching<sup>34</sup> and in consistent-mapping visual search paradigms.<sup>35–37</sup> Although older adults did not show greater improvement when working-memory load was high,<sup>34</sup> nor did they show as much stimulus-specific learning as did younger participants,<sup>36</sup> older adults retained what they learned up to 16 months.<sup>37</sup> In the present study, older pilots conceivably improved their reactions to oncoming traffic by learning task switching and visual search strategies helpful to performance. Greater improvement among older adults has rarely been reported in longitudinal studies employing paper-and-pencil neuropsychological or intellectual ability tests,<sup>31,32</sup> but see reference<sup>38</sup>. Our atypical finding will need to be replicated in the independent cohort of aviators we are presently enrolling.

Some cross-sectional studies found that aviation expertise moderated age differences in pilot performance.<sup>8,39,40</sup> Other studies found that while aviation expertise significantly helped performance, expertise did not significantly moderate the influence of age.<sup>10,21</sup> The present study did not find that aviation expertise moderated the impact of age on longitudinal flight simulator performance. Because the numbers of participants in the extremes of the age and expertise distributions were modest, statistical power for the test of an age  $\times$  expertise interaction was less than it would be in an extreme-groups design.<sup>41</sup> Clearly, a longer duration of follow-up is crucial to examining an age-moderating effect of expertise on specialized skill domains.<sup>42</sup>

These findings have broader implications beyond aviation to the general issue of aging in the workplace. Several issues emerge from an aging workforce, including technological developments, training and retraining, retirement, physical capacity, health, and performance.<sup>43</sup> Middle-aged workers, for example, can be retrained as effectively as young workers, while older workers also can be retrained but less efficiently than their younger counterparts.<sup>44</sup> If retirement ages become increasingly delayed, objective assessments of workplace competence will become essential for older workers, especially when the occupation is viewed as a public safety concern. On the one hand, there is rising incidence of medical and neurologic problems with age.<sup>45</sup> On the other hand, older expert workers may be able to adapt to normal age-associated changes through increased reliance on domain-specific knowledge and procedural memories, which are less age-sensitive, and by adopting strategies that help maintain successful performance and minimize errors.<sup>4,5,44,46,47</sup> In order to fairly and objectively assess occupational competency, it is necessary to incorporate measures rich for domain-relevant knowledge and strategies.<sup>8,43,48</sup>

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

The authors thank Helena Kraemer, PhD, for biostatistical consulting and Katy Castile, Tiffany Doelger, and Anne Lademan for recruiting and testing participants. They also thank the aviator study participants for their donation of time and for being inspirational role models of intellectual exploration.

Supported by the Sierra-Pacific Mental Illness Research, Education, and Clinical Center (MIRECC) and the Medical Research Service of the Department of Veterans Affairs, and by NIA grants P30 AG 17824 and R37 AG 12713 (with a supplement for underrepresented minorities to Dr. Kennedy).

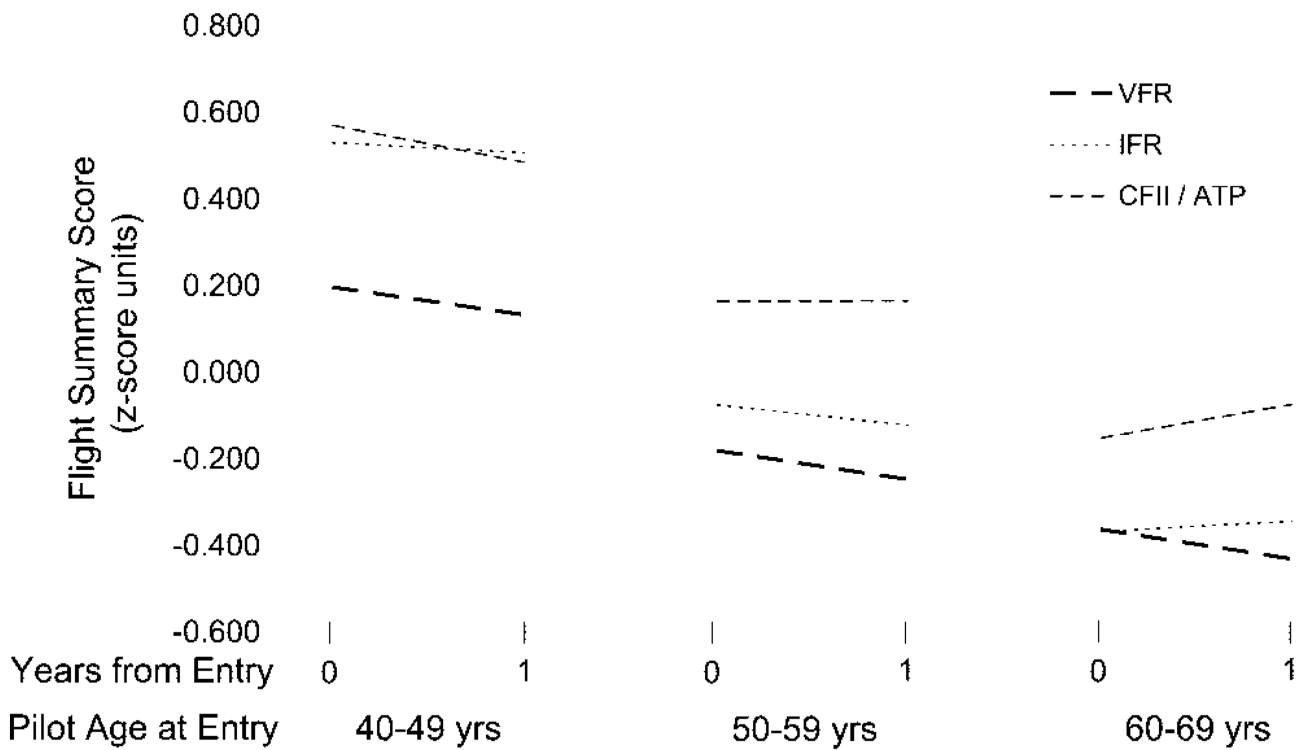


## References

1. Wilkening R. The age 60 rule: age discrimination in commercial aviation. *Aviat Space Environ Med* 2002;73:194–202. [PubMed: 11908884]
2. Stuck AE, van Gorp WG, Josephson KR, Morgenstern H, Beck JC. Multidimensional risk assessment versus age as criterion for retirement of airline pilots. *J Am Geriatr Soc* 1992;40:526–532. [PubMed: 1634710]
3. Hoyer WJ, Ingolfsdottir D. Age, skill, and contextual cuing in target detection. *Psychol Aging* 2003;18:210–218. [PubMed: 12825771]
4. Krampe RT. Aging, expertise and fine motor movement. *Neurosci Biobehav Rev* 2002;26:769–776. [PubMed: 12470688]
5. Krampe RT, Ericsson KA. Maintaining excellence: deliberate practice and elite performance in young and older pianists. *J Exp Psychol Gen* 1996;125:331–359. [PubMed: 8945787]
6. Masunaga H, Horn J. Expertise and age-related changes in components of intelligence. *Psychol Aging* 2001;16:293–311. [PubMed: 11405317]
7. Meinz EJ. Experience-based attenuation of age-related differences in music cognition tasks. *Psychol Aging* 2000;15:297–312. [PubMed: 10879584]
8. Morrow DG, Ridolfo HE, Menard WE, et al. Environmental support promotes expertise-based mitigation of age differences on pilot communication tasks. *Psychol Aging* 2003;18:268–284. [PubMed: 12825776]
9. Yesavage JA, Taylor JL, Mumenthaler MS, Noda A, O'Hara R. Relationship of age and simulated flight performance. *J Am Geriatr Soc* 1999;47:819–823. [PubMed: 10404925]
10. Taylor JL, O'Hara R, Mumenthaler MS, Rosen AC, Yesavage JA. Cognitive ability, expertise, and age differences in following air-traffic control instructions. *Psychol Aging* 2005;20:117–133. [PubMed: 15769218]
11. Hofer SM, Sliwinski MJ. Understanding ageing. An evaluation of research designs for assessing the interdependence of ageing-related changes. *Gerontology* 2001;47:341–352. [PubMed: 11721149]
12. Kraemer HC, Yesavage JA, Taylor JL, Kupfer D. How can we learn about developmental processes from cross-sectional studies, or can we? *Am J Psychiatry* 2000;157:163–171. [PubMed: 10671382]
13. Kay, GG. *CogScreen Aeromedical edition professional manual*. Odessa, FL: Psychological Assessment Resources; 1995.
14. Salthouse, TA. *Mechanisms of age-cognition relations in adulthood*. Hillsdale, NJ: Lawrence Erlbaum Associates; 1992.
15. Taylor JL, O'Hara R, Mumenthaler MS, Yesavage JA. Relationship of CogScreen-AE to flight simulator performance and pilot age. *Aviat Space Environ Med* 2000;71:373–380. [PubMed: 10766461]
16. Yesavage JA, Mumenthaler MS, Taylor JL, et al. Donepezil and flight simulator performance: effects on retention of complex skills. *Neurology* 2002;59:123–125. [PubMed: 12105320]
17. Kraemer HC, Blasey CM. Centring in regression analyses: a strategy to prevent errors in statistical inference. *Int J Methods Psychiatr Res* 2004;13:141–151. [PubMed: 15297898]
18. Li G, Baker SP, Qiang Y, Grabowski JG, McCarthy ML. Driving-while-intoxicated history as a risk marker for general aviation pilots. *Accid Anal Prev* 2005;37:179–184. [PubMed: 15607289]
19. Li G, Baker SP, Grabowski JG, Qiang Y, McCarthy ML, Rebok GW. Age, flight experience, and risk of crash involvement in a cohort of professional pilots. *Am J Epidemiol* 2003;157:874–880. [PubMed: 12746239]
20. Bellenkes AH, Wickens CD, Kramer AF. Visual scanning and pilot expertise: the role of attentional flexibility and mental model development. *Aviat Space Environ Med* 1997;68:569–579. [PubMed: 9215461]
21. Morrow DG, Miller LM, Ridolfo HE, Menard W, Stine-Morrow EA, Magnor C. Environmental support for older and younger pilots' comprehension of air traffic control information. *J Gerontol B Psychol Sci Soc Sci* 2005;60:P11–P18. [PubMed: 15643033]

22. Braune, R.; Wickens, CD. Final Technical Report EPL-83-7/NAMRL-83-2. Champaign, IL: University of Illinois, Engineering Psychology Laboratory, Department of Psychology; 1984. Individual differences and age-related performance assessment in aviators. Part 2: Initial battery validation.
23. Wiggins M, O'Hare D. Expertise in aeronautical weather-related decision making: a cross-sectional analysis of general aviation pilots. *J Exp Psychol Appl* 1995;1:305–320.
24. Ericsson KW, Lehmann AC. Expert and exceptional performance: evidence of maximal adaptation to task constraints. *Annu Rev Psychol* 1996;47:273–305. [PubMed: 15012483]
25. Draganski B, Gaser C, Busch V, Schuierer G, Bogdahn U, May A. Neuroplasticity: changes in grey matter induced by training. *Nature* 2004;427:311–312. [PubMed: 14737157]
26. Jastrzembski TS, Charness N, Vasyukova C. Expertise and age effects on knowledge activation in chess. *Psychol Aging* 2006;21:401–405. [PubMed: 16768584]
27. Peres M, Van De Moortele PF, Pierard C, et al. Functional magnetic resonance imaging of mental strategy in a simulated aviation performance task. *Aviat Space Environ Med* 2000;71:1218–1231. [PubMed: 11439722]
28. Morrow D, Yesavage J, Leirer V, Tinklenberg J. Influence of aging and practice on piloting tasks. *Exp Aging Res* 1993;19:53–70. [PubMed: 8444267]
29. Taylor JL, Dolhert N, Morrow D, Friedman L, Yesavage JA. Acute and 8-hour effects of alcohol (0.08% BAC) on younger and older aircraft pilots' simulator performance. *Aviat Space Environ Med* 1994;65:718–725. [PubMed: 7980331]
30. Rabbitt P, Diggle P, Smith D, Holland F, Mc Innes L. Identifying and separating the effects of practice and of cognitive ageing during a large longitudinal study of elderly community residents. *Neuropsychologia* 2001;39:532–543. [PubMed: 11254936]
31. Ronnlund M, Nyberg L, Backman L, Nilsson LG. Stability, growth, and decline in adult life span development of declarative memory: cross-sectional and longitudinal data from a population-based study. *Psychol Aging* 2005;20:3–18. [PubMed: 15769210]
32. Salthouse TA, Schroeder DH, Ferrer E. Estimating retest effects in longitudinal assessments of cognitive functioning in adults between 18 and 60 years of age. *Dev Psychol* 2004;40:813–822. [PubMed: 15355168]
33. Wilson RS, Beckett LA, Barnes LL, et al. Individual differences in rates of change in cognitive abilities of older persons. *Psychol Aging* 2002;17:179–193. [PubMed: 12061405]
34. Kramer AF, Hahn S, Gopher D. Task coordination and aging: explorations of executive control processes in the task switching paradigm. *Acta Psychol (Amst)* 1999;101:339–378. [PubMed: 10344190]
35. Plude DJ, Hoyer WJ. Adult age differences in visual search as a function of stimulus mapping and processing load. *J Gerontol* 1981;36:598–604. [PubMed: 7264245]
36. Rogers WA. Age differences in visual search: target and distractor learning. *Psychol Aging* 1992;7:526–535. [PubMed: 1466821]
37. Fisk AD, Hertzog C, Lee MD, Rogers WA, Anderson-Garlach M. Long-term retention of skilled visual search do young adults retain more than old adults? *Psychol Aging* 1994;9:206–215. [PubMed: 8054168]
38. Rabbitt P, Diggle P, Holland F, McInnes L. Practice and drop-out effects during a 17-year longitudinal study of cognitive aging. *J Gerontol B Psychol Sci Soc Sci* 2004;59:P84–P97. [PubMed: 15014091]
39. Morrow D, Leirer V, Altieri P, Fitzsimmons C. When expertise reduces age differences in performance. *Psychol Aging* 1994;9:134–148. [PubMed: 8185861]
40. Tsang PS, Shaner TL. Age, attention, expertise, and time-sharing performance. *Psychol Aging* 1998;13:323–347. [PubMed: 9640591]
41. McClelland GH, Judd CM. Statistical difficulties of detecting interactions and moderator effects. *Psychol Bull* 1993;114:376–390. [PubMed: 8416037]
42. Salthouse T. Mental exercise and mental aging: evaluating the validity of the “use it or lose it” hypothesis. *Perspectives Psychol Sci* 2006;1:68–87.
43. Sterns HL, Miklos SM. The aging worker in a changing environment: organizational and individual issues. *J Vocat Behav* 1995;47:248–268.

44. Charness N, Kelley CL, Bosman EA, Mottram M. Word-processing training and retraining: effects of adult age, experience, and interface. *Psychol Aging* 2001;16:110–127. [PubMed: 11302360]
45. Miech RA, Breitner JC, Zandi PP, Khachaturian AS, Anthony JC, Mayer L. Incidence of AD may decline in the early 90s for men, later for women: the Cache County study. *Neurology* 2002;58:209–218. [PubMed: 11805246]
46. Morrow, D.; Leirer, V. Fisk, AD.; Rogers, WA. *Handbook of human factors and the older adult*. San Diego: Academic Press; 1997. Aging, pilot performance, and expertise; p. 199-230.
47. Czaja, SJ. Technological change and the older worker. In: Birren, JE.; Schaie, KW., editors. *Handbook of the psychology of aging*. 5th ed.. San Diego: Academic Press; 2001. p. 547-568.
48. Vicente KJ, Wang JH. An ecological theory of expertise effects in memory recall. *Psychol Rev* 1998;105:33–57. [PubMed: 9450371]



**Figure.**

Patterns of longitudinal change in flight simulator performance by age and expertise group. Plotted in the figure are the mean intercepts at the age of entry and the mean slopes for each group. Time 0 is the estimated baseline level (i.e., the intercept). Time 1 is the estimated change from baseline 1 year later (intercept + slope). We note that the numbers of participants in the extremes of the age and expertise distributions were modest (5 to 9). Among participants who were 40 to 49 years of age at study entry, 9 were VFR-rated, 6 were IFR-rated, and 5 had attained CFII or ATP certifications. Of those who were 50 to 59 years of age at study entry, 17 were VFR-rated, 29 were IFR-rated, and 15 had attained CFII or ATP certifications. Of those 60 to 69 years of age at study entry, 6 were VFR-rated, 25 were IFR-rated, and 6 had attained CFII or ATP certifications.

**Table 1**

Demographic and cognitive ability characteristics (mean  $\pm$  SD) at study entry of the 118 longitudinally followed participants, grouped according to pilot license rating, the study's measure of aviation expertise\*

	VFR (n = 32)	IFR (n = 60)	CFII/ATP (n = 26)
Age, y, mean $\pm$ SD <sup>†</sup>	54.3 $\pm$ 6.5	58.4 $\pm$ 6.7	55.1 $\pm$ 6.0
Education, y, mean $\pm$ SD	16.5 $\pm$ 2.3	16.9 $\pm$ 1.9	16.7 $\pm$ 2.0
Percent women	19	23	8
Percent white, non-Hispanic	91	98	92
Total flight time, h, mean $\pm$ SD <sup>‡</sup>	1,056 $\pm$ 1,393	1,682 $\pm$ 1,625	5,155 $\pm$ 2,932
Recent flight time, h past m, mean $\pm$ SD <sup>§</sup>	5.8 $\pm$ 8.0	7.5 $\pm$ 6.3	16.1 $\pm$ 17.4
Percent medical Class I, II, or III <sup>  </sup>	0, 19, 81	2, 38, 60	31, 61, 8
Percent medical waivers <sup>¶</sup>	6.3	10.0	3.9
Cognitive variables, mean $\pm$ SD <sup>**</sup>			
Dual task: dual tracking error	70.0 $\pm$ 26.2	71.5 $\pm$ 22.8	68.9 $\pm$ 22.3
Manikin throughput	34.1 $\pm$ 9.6	30.0 $\pm$ 8.5	34.4 $\pm$ 9.0
Pathfinder combined throughput	50.4 $\pm$ 12.5	46.1 $\pm$ 15.7	49.7 $\pm$ 13.4
Shifting attention instruction throughput	76.1 $\pm$ 12.1	66.8 $\pm$ 15.4	69.2 $\pm$ 12.2
Speed of processing z-score composite	0.3 $\pm$ 0.8	-0.1 $\pm$ 0.9	-0.2 $\pm$ 0.9

\* VFR: rated for flying under visual conditions, which restricts a pilot to flying only in good visibility conditions, is the rating given to pilots when they first obtain a license. IFR: instrument rated, which allows a pilot to fly in poorer visibility conditions using navigational instruments. An IFR rating requires at least 40 hours of instrument time, where pilots learn how to use instruments and radar information to achieve precise navigation and maneuvering and learn more about air traffic control instructions and procedures. CFII: certified flight instructor of pilots in training for IFR. ATP: certified to fly air-transport planes. Major airline captains have the ATP rating.

<sup>†</sup> IFR group was older than the VFR and CFII/ATP groups ( $p < 0.05$ ).

<sup>‡</sup> CFII/ATP group had more total flight hours than the IFR group, and the IFR group had more total flight hours than the VFR group ( $p < 0.05$ ).

<sup>§</sup> The CFII/ATP group had more recent flight hours than the VFR group ( $p < 0.05$ ).

<sup>||</sup> The majority of VFR and IFR pilots had Class III medical certificates, whereas the majority of CFII/ATP pilots had Class II certificates ( $p < 0.0001$ ). Pilots are required to pass periodic medical examinations in order to fly. Class I and II medical certificates have higher standards for distant and near vision than Class III; a Class I certificate requires an ECG at age 35 and annually after age 40, but Class II and III do not routinely require them. A Class I medical certificate indicates that the pilot has passed a medical examination within the past 6 months; a Class II medical certificate indicates passing the examination within the past 1 year; a Class III medical certificate indicates passage within the past 2 years. Due to this medical monitoring, pilots tend to be in good health.

<sup>¶</sup> A few participants (9/118) had a Federal Aviation Administration (FAA) medical waiver, which allows a pilot to have a valid FAA medical certificate despite having an otherwise disqualifying medical condition.

\*\* See tables E-1 and E-2 on the *Neurology* Web site at [www.neurology.org](http://www.neurology.org) for more information regarding these variables.

Relations of age and aviation expertise to baseline level and average annual rate of change in flight simulator performance: Summary of general linear model results (parameter estimates  $\pm$  SE and overall  $R^2$ )

**Table 2**

Flight measure	Dependent variable*	Model terms						Overall model $R^2$		
		Intercept ( $\beta_0$ )		Age ( $\beta_1$ )		Expertise ( $\beta_2$ )			Age $\times$ expertise ( $\beta_3$ )	
		Estimate	SE	Estimate	SE	Estimate	SE		Estimate	SE
Summary score	Baseline level	-0.046	0.041	-0.038 <sup>§</sup>	0.006	0.155 <sup>†</sup>	0.060	0.000	0.009	0.28
	Rate of change	-0.025 <sup>†</sup>	0.011	0.004 <sup>‡</sup>	0.002	0.039 <sup>†</sup>	0.015	0.004	0.002	0.11
Communication	Baseline level	-0.087	0.058	-0.037 <sup>§</sup>	0.009	0.197 <sup>†</sup>	0.085	-0.006	0.013	0.17
	Rate of change	-0.091 <sup>§</sup>	0.014	0.000	0.002	0.045 <sup>†</sup>	0.020	0.001	0.003	0.04
Traffic avoidance	Baseline level	-0.036	0.046	-0.044 <sup>§</sup>	0.007	0.107	0.067	-0.010	0.011	0.28
	Rate of change	0.041 <sup>†</sup>	0.020	0.008 <sup>‡</sup>	0.003	0.023	0.030	0.006	0.005	0.07
Emergencies	Baseline level	-0.031	0.078	-0.029 <sup>†</sup>	0.012	0.110	0.114	0.009	0.018	0.06
	Rate of change	0.019	0.021	0.001	0.003	0.015	0.031	0.003	0.005	0.00
Approach	Baseline level	-0.028	0.055	-0.042 <sup>§</sup>	0.008	0.206 <sup>†</sup>	0.081	0.008	0.013	0.22
	Rate of change	-0.072 <sup>‡</sup>	0.025	0.008	0.004	0.071	0.037	0.004	0.006	0.06

\* Baseline level and average annual rate of change variables were estimated by regressing each participant's scores on age at test.

<sup>†</sup>  $p < 0.05$ .

<sup>‡</sup>  $p < 0.01$ .

<sup>§</sup>  $p < 0.0001$ .