

# UCSF

## UC San Francisco Previously Published Works

### Title

Pilot Expertise and Hippocampal Size: Associations with Longitudinal Flight Simulator Performance

### Permalink

<https://escholarship.org/uc/item/6rj4z7fv>

### Journal

Aerospace Medicine and Human Performance, 83(9)

### ISSN

2375-6314

### Authors

Adamson, Maheen M  
Bayley, Peter J  
Scanlon, Blake K  
[et al.](#)

### Publication Date

2012-09-01

### DOI

10.3357/ asem.3215.2012

Peer reviewed



Published in final edited form as:

*Aviat Space Environ Med.* 2012 September ; 83(9): 850–857. doi:10.3357/asem.3215.2012.

## Pilot Expertise and Hippocampal Size: Associations with Longitudinal Flight Simulator Performance

MAHEEN M. ADAMSON, PETER J. BAYLEY, BLAKE K. SCANLON, MICHELLE E. FARRELL, BEATRIZ HERNANDEZ, MICHAEL W. WEINER, JEROME A. YESAVAGE, JOY L. TAYLOR

War Related Illness and Injury Study Center (WRIISC) and the Sierra-Pacific MIRECC, Veterans Affairs, Palo Alto Health Care System, Palo Alto, CA; Psychiatry and Behavioral Sciences, Stanford School of Medicine, Stanford, CA; the University of California, San Francisco, CA; and the DVA Medical Center, San Francisco, CA.

### Abstract

**Background**—Previous research suggests that the size of the hippocampus can vary in response to intensive training (e.g., during the acquisition of expert knowledge). However, the role of the hippocampus in maintenance of skilled performance is not well understood. The Stanford/Veterans Affairs Aviation MRI Study offers a unique opportunity to observe the interaction of brain structure and multiple levels of expertise on longitudinal flight simulator performance.

**Methods**—The current study examined the relationship between hippocampal volume and three levels of aviation expertise, defined by pilot proficiency ratings issued by the U.S. Federal Aviation Administration (11). At 3 annual time points, 60 pilots who varied in their level of aviation expertise (ages ranging from 45 to 69 yr) were tested.

**Results**—At baseline, higher expertise was associated with better flight simulator performance, but not with hippocampal volume. Longitudinally, there was an Expertise  $\times$  Hippocampal volume interaction, in the direction that a larger hippocampus was associated with better performance at higher levels of expertise.

**Discussion**—These results are consistent with the notion that expertise in a cognitively demanding domain involves the interplay of acquired knowledge (‘mental schemas’) and basic hippocampal-dependent processes.

### Keywords

hippocampus; expertise; skill; aviation; memory

---

THE HIPPOCAMPUS IS part of a system of anatomically related structures in the medial temporal lobe (MTL) that supports the capacity for long-term memory (24). The MTL is

---

Address correspondence and reprint requests to: Maheen M. Adamson, Ph.D., War Related Illness and Injury Study Center (WRIISC), Veterans Affairs, Palo Alto Health Care System, 3801 Miranda Avenue (151Y), Palo Alto, CA, 94304; madamson@stanford.edu.  
*Authors and affiliations:* Maheen M. Adamson, Ph.D., Peter J. Bayley, Ph.D., Blake K. Scanlon, Ph.D., Michelle E. Farrell, B.A., Jerome A. Yesavage, M.D., and Joy L. Taylor, Ph.D., Veterans Affairs, Palo Alto Health Care System, Palo Alto, CA; Beatriz Hernandez, B.A., Psychiatry and Behavioral Sciences, Stanford Medical School, Stanford, CA; and Michael W. Weiner, M.D., University of California, San Francisco, and the DVA Medical Center San Francisco, MRS Unit, VA Medical Center, San Francisco, CA.

critical for learning (i.e., encoding) and retrieval of long-term memory for facts and events (24). As a whole, the hippocampus and surrounding MTL are vital for this “declarative” or “explicit” form of memory, but is not thought to be involved in other forms of “nondeclarative” or “implicit” long-term memory such as habit learning or simple forms of conditioning (24). In short, the hippocampus helps encode and retain conscious experiences.

It has been suggested that hippocampal volume can vary as a result of intensive training such as during acquisition of expert knowledge (17). Evidence for differences in hippocampal volume associated with acquired expertise has come primarily from cross-sectional studies comparing novices and experts. For example, studies have reported differences in hippocampal volume between taxi drivers and controls (21). These findings have been interpreted in terms of the critical role of the hippocampus in spatial memory and navigation. In addition, several longitudinal studies have suggested a more causative link between structural and functional changes in the hippocampus and extensive training in domains such as music (16), architecture (18), persuasive communication (19), and juggling (5). To date, only one functional neuroimaging study of expertise acquisition has been reported: after 2 wk of intensive music training, an enhancement in hippocampal activation to the temporal novelty of sounds was found (16). Taken together, these studies suggest that measurable changes may occur in the hippocampus as a result of intensive training.

In the present study, we investigated the role of hippocampal volume in maintaining skilled performance. Although we examined the association between hippocampal volume and expertise, as has been done in previous studies such as in taxi drivers, our main aim was to examine the roles of hippocampal volume and aviation expertise in maintaining skilled performance as measured by annual flight simulator testing. We characterized aviation expertise in terms of the proficiency ratings that pilots acquire through formal training and certification (26). This measure of expertise is akin to the construct of deliberate practice, which emphasizes intensive engagement in activities that are designed to master specific goals in a particular skill (10).

Subjects were enrollees of the Stanford/Veterans Affairs (VA) Aviation MRI Study, which is also examining the influences of the apolipoprotein E (APOE) gene, the major genetic susceptibility factor for late-onset Alzheimer’s disease (AD). The  $\epsilon 4$  allele of APOE is associated with increased risk for the development of AD dementia at an earlier age of onset (1). Other isoforms of APOE are considered to be neutral (APOE  $\epsilon 3$ ) or protective (APOE  $\epsilon 2$ ) for developing AD. APOE  $\epsilon 4$  has also been associated with a reduced hippocampal volume in some longitudinal studies of cognitively normal older adults and in patients with early AD, but not in our sample of pilots (3).

Expert motor skills are typically associated with brain regions, including the striatum, cerebellum, basal ganglia, prefrontal cortex, and parietal lobes (3). For instance, Adamson et al. (3) observed a significant interaction between expertise and parietal cortex volume, such that parietal gray matter was positively associated with performance in expert aviators. This study reported on the relationship between aviation expertise and brain volume on baseline flight simulator performance (3). The results suggest that brain structure plus knowledge interact to support skilled performance. To systematically tease apart the major independent

variables and moderator effects in the present study, we follow the procedures outlined in Adamson et al. (3). In essence, moderators help identify in whom—in this case, experts or novices—brain structure is more important in subserving performance.

In assessing longitudinal flight simulator performance, we strove to minimize practice effects by familiarizing pilots with the flight simulator scenario prior to their annual testing. Performance was highest at baseline, on average, and then declined over time (25). We tested the main effects as well as the interaction of expertise and hippocampal size in mitigating decline in complex skilled performance. In secondary analyses we tested main effects and interactions between expertise and volumes of other brain structures, including the cerebellum, basal ganglia, prefrontal, and parietal cortex. Thus, the main hypothesis tested in the present study was: aviation expertise will moderate the impact of hippocampal volume on change in flight simulator performance measured across three annual tests.

## METHODS

### Subjects

A total of 60 general aviation pilots (11 women and 49 men; 55 white) were studied (see Table I for demographics according to the three aviation expertise levels). Enrollment criteria for MRI study inclusion were: age 45 or older, current FAA medical certificate, and currently flying. Each participant was classified into one of three levels of aviation expertise (least, moderate, and most expertise). The three levels of expertise were defined in terms of the FAA-issued pilot proficiency ratings that pilots had attained prior to entering the study: 1) least expert (VFR – rated for flying under visual conditions, which restricts a pilot to flying only in good visibility conditions. VFR is the rating given to pilots when they first obtain a license.); 2) moderate expertise (IFR – instrument rated, which allows a pilot to fly in poorer visibility conditions using navigational instruments); and 3) most expert (CFII/ATP – pilots were certified flight instructors of pilots in training for IFR, or they had also become certified to fly air transport planes. Major airline captains and other professional pilots have the ATP rating). Each rating requires progressively more advanced training and more hours of flight experience (26).

MRI study subjects were selectively recruited from the ongoing longitudinal Stanford/VA Aviation Study so that approximately 50% of MRI subjects were APOE  $\epsilon 4$  carriers ( $\epsilon 3/4$  or  $\epsilon 4/4$ ) and 50% were  $\epsilon 3/3$  [for details on genotyping methods, see Adamson et al. (2)]. Informed consent, approved by the Stanford University and VA Palo Alto Health Care System Institutional Review Boards, was obtained from each participant. At baseline and each annual follow-up visit, pilots reported total number of flight hours accumulated to date.

### Equipment

Pilots “flew” in a Frasca 141 flight simulator (Urbana, IL). The simulator was linked to a computer specialized for graphics (Dell Precision Workstation with a Fedora Linux Operating system, OpenGL C++ software specially customized for a high-end NVIDIA graphics processor) that generated a “through-the-window” visual environment and continuously collected data concerning the aircraft’s position and communication

frequencies (see <http://www.stanford.edu/people/yesavage/AIR.html>). 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, subjects had six practice flights in the simulator to gain familiarity with the flight scenario used throughout the study. Subjects typically completed their practice flights during a 1- to 3-wk period, after which they had a 3-wk break before returning for the baseline visit. At the baseline and the two annual follow-up visits, subjects 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 tests. The entire test day lasted approximately 6 h, including a 40- to 60-min lunch break. The flight scenario was based on our previous studies (3,26). 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 baseline visit mean and SD of the 141 Aviation Study subjects enrolled during 1996 – 2001 (scores on the morning and afternoon flights were averaged). Flight simulator performance was measured as a single composite z-score (28).

MRI data were acquired on a 1.5-Tesla (GE Medical Systems, Milwaukee, WI) MRI scanner. A 3D spoiled GRASS MRI of the entire brain was acquired using a standard head coil, TR/TE = 9/2 ms, 15° flip angle, perpendicular to the long axis of the hippocampus [ $1.00 \times 1.00 \text{ mm}^2$  in plane resolution, 1.5-mm coronal slices, no skip; for details, see Adamson et al.(3)]. The average time lag between the baseline simulator visit and the MRI scan was 3.85 yr (SD = 3.0 yr) because funding for MRIs was obtained after the longitudinal simulator study was underway. The time lag was not significantly correlated with other demographic variables, including age at the baseline simulator visit ( $r = -0.18$ ,  $P = 0.17$ ). Also, there was no significant correlation of the time lag with the primary outcome variable, longitudinal change in flight simulator performance ( $r = 0.02$ ,  $P = 0.87$ ).

MRI analyses were performed by Dr. Weiner's lab at UCSF. Cortical reconstruction and volumetric segmentation were performed with the Freesurfer image analysis suite (v. 4.5, <http://surfer.nmr.mgh.harvard.edu/>). The technical details of Freesurfer are described in prior publications (12). Hippocampal voluming was done using a commercial high-dimensional brain-mapping tool [Medtronic Surgical Navigation Technologies (SNT), Louisville, CO; for details of these procedures, see Adamson et al. (3)]. Total intracranial volume (TIV) equaled the sum of all segmented tissue inside the skull (excluding cerebellum and including fluid). TIV ( $1711.8 \pm 177.3 \text{ cc}$ ) correlated strongly with whole brain volume ( $r = 0.80$ ) and was used for normalization of hippocampus, whole brain, and lobar volumes.

## Statistical Analysis

The primary longitudinal outcome was average rate of change per year (slope score) in overall flight simulator performance (28). Slope scores were computed separately for each participant by regressing the three flight summary scores on the subject's age at each annual simulator test. All analyses were performed in SAS (Version 9.1.3, Cary, NC). To test the main hypothesis that aviation expertise moderates the impact of hippocampal volume on longitudinal flight simulator performance, a multiple regression model was fitted to test the main effects of hippocampal volume, expertise, and their interaction (SAS PROC GLM). Other terms in the GLM were the intercept estimate, age, baseline flight simulator performance, APOE  $\epsilon 4$  carrier status, and Expertise  $\times$  APOE  $\epsilon 4$  carrier status in line with our previous work on APOE  $\epsilon 4$  (2,25,26). Expertise was centered [-1 for least expertise (VFR), 0 for moderate expertise, 1 for most expertise (CFII/ATP)] and APOE  $\epsilon 4$  was coded as -0.5 or  $\pm$  0.5. Age was a continuous covariate centered at the median age at time of baseline simulator visit (56.85 yr).

The influence of potentially confounding demographic variables and a few outlying flight scores were of concern. As reported in the Results, aviation expertise and education were positively correlated, even though years of education beyond 20 yr were truncated at 20. Also, there were no women in the expert group. Therefore, education and gender were included in a stepwise regression model. To address outliers, we truncated any flight component scores that were more than 3 SDs from the group mean and recalculated the baseline and slope scores (3). The stepwise analysis reported used these truncated scores. All calculated *P*-values were two-sided. Some of these correlations among age, aviation expertise, education, brain volumes, and baseline performance measures have been reported previously on fewer numbers of subjects (2,3).

## RESULTS

Table I summarizes the characteristics of the subjects at entry grouped by expertise level. In terms of expertise, there were differences among the groups in hours of flight experience, age, and years of education. Specifically, each successive level of expertise was associated with more hours of total flight experience, as would be expected, overall [ $F(2, 57) = 23.86, P < 0.0001$ ]. There was an association between education and aviation expertise [ $F(2, 57) = 3.33, P = 0.043$ ]. Apart from hours of flight experience and education, there were no significant differences among the expertise groups in the remaining demographic measures (see Table I for means and SDs).

In order for aviation expertise to moderate hippocampal size on rate-of-change in flight simulator performance, expertise and hippocampal volume should not correlate. This correlation was not significant ( $r = 0.09, P = 0.50$ ), indicating that pilots with the most expertise did not have larger hippocampi than pilots with lesser expertise. Also, expertise did not diminish the impact of age on hippocampal volume (Expertise  $\times$  Age interaction:  $F < 1$ ).

With this condition satisfied, we proceeded to test the main hypothesis by conducting the multiple regression analysis described in the Methods. There were no significant main effects of expertise [ $F(1, 52) = 2.72, P = 0.105$ ], age [ $F(1, 52) = 2.30, P = 0.136$ ], or APOE

$\epsilon 4$  status [ $F < 1$ ]; nor was there a significant main effect of hippocampal volume [ $F(1, 52) = 1.43, P = 0.238$ ] on average rate of change per year (slope score) in overall flight simulator performance. There was a significant main effect of baseline flight simulator performance [ $F(1, 52) = 4.95, P = 0.031$ ].

The predicted interaction between expertise and hippocampal volume was observed [ $F(1, 52) = 6.48, P = 0.014, ES = 0.33$  (see Table II for  $\beta$  parameter estimates)]. Post hoc multiple regression analyses showed that the interaction was primarily driven by differences between moderate vs. least expertise pilots ( $P < 0.05$ ). The Expertise  $\times$  APOE  $\epsilon 4$  interaction approached significance [ $F(1, 52) = 3.84, P = 0.055, ES = 0.25$ ].<sup>1</sup>

Fig. 1 provides a scatter plot of rate of change in simulator performance and hippocampal volume in the three expertise groups. The most expert pilots tended to have less longitudinal decline in flight simulator performance (average rate of decline =  $-0.019 \pm 0.12$  per year); within the expert group larger hippocampal volume was associated with better performance over time, although this was not significant ( $r = 0.35, P > 0.1$ ). This is in contrast to the least expert pilots, who tended to have a steeper decline in performance over time (average rate of decline =  $-0.067 \pm 0.23$  per year) and a nonsignificant negative relationship with hippocampal volume ( $r = -0.35, P > 0.1$ ). Pilots with intermediate expertise were in between the two expertise levels with no apparent relationship with hippocampal volume ( $r = -0.16, P > 0.1$ ).

In light of the overlapping influences of aviation expertise and accumulated hours, we performed a stepwise regression in which all main effects, interactions of interest (Hippocampal volume  $\times$  Expertise and Expertise  $\times$  APOE  $\epsilon 4$ ), accumulated hours, and potentially confounding demographic variables were entered into the model (Table III). There were two significant predictors. The first was Expertise  $\times$  Hippocampal volume [ $F(2, 57) = 5.27; P = 0.03; \text{partial } R^2 = 0.074$ ], followed by Expertise  $\times$  APOE  $\epsilon 4$  [ $F(2, 57) = 4.80; P = 0.033; \text{partial } R^2 = 0.072$ ]. Therefore, longitudinal flight simulator performance was predicted by the Expertise  $\times$  Hippocampal volume interaction. None of the other brain regions (or their interaction with Expertise) predicted longitudinal flight simulator performance ( $P > 0.1$ ).

Analyses of component scores for longitudinal simulator performance revealed a significant Expertise  $\times$  Hippocampal volume interaction ( $\beta = 0.378, SE = 0.160, P = 0.022$ ) and a significant Expertise  $\times$  APOE  $\epsilon 4$  interaction ( $\beta = 0.293, SE = 0.133, P = 0.032$ ) in the approach component of the flight. Traffic avoidance showed the strongest main effect of hippocampal volume ( $\beta = -0.171, SE = 0.082, P = 0.042$ ), such that smaller hippocampal volume was associated with improved performance with repeated testing.

## DISCUSSION

We sought to identify the contributions of aviation expertise and hippocampal volume on change in flight simulator performance measured across three annual tests. We did not find that expert pilots had larger hippocampal volumes on average, although previous studies have reported increases in hippocampal subvolumes and gray matter density associated with



intensive training or expert knowledge. As hypothesized, we found that aviation expertise moderated the impact of hippocampal volume on change in flight simulator performance. Specifically, the Expertise  $\times$  Hippocampal volume interaction was in the direction that a larger hippocampus was associated with better maintenance of performance at a higher level of expertise. Finally, larger hippocampal volume was not associated with better maintenance of performance in the sample as a whole, as there was not a main effect of hippocampal volume on longitudinal performance. Taken together, these findings suggest complex relationships among hippocampal volume, expertise, and performance.

The hippocampus has a well-established role in the formation and consolidation of memory and our finding emphasizes the importance of memory processes in the flight simulator task. The hippocampus is critical for the formation of declarative memory, which refers to the conscious recollection of both facts and events (24). Declarative memory includes many memory processes, including associative memory, in which previously unrelated information is bound together (15), and spatial memory (21). The hippocampus is also critical for the consolidation of new declarative memories into long-term storage. Through a gradual process of consolidation and reorganization, connections among cortical regions are thought to be progressively strengthened until the cortical memory eventually becomes independent of the hippocampus. The importance of memory processes in the flight simulator task is also consistent with our earlier cross-sectional findings (3) that simulator performance was best among expert pilots, especially expert pilots with larger parietal lobe volume. The parietal lobe has been implicated in many aspects of declarative memory, including spatial working memory (20), episodic memory and retrieval (6), and consolidation into long-term memory (22).

How might the hippocampus be related to the maintenance of flight simulator performance in our group of expert pilots? One obvious answer is that the declarative memory functions supported by the hippocampus are critical for acquiring and/or maintaining expert flying skills. Cross-sectional studies have identified a relationship between expertise and hippocampal volume. Maguire and coworkers have extensively documented posterior hippocampal enlargement in London taxi drivers (21). These results were attributed to processes involved in learning, representing and using spatial representations of a highly complex and large-scale environment. Another cross-sectional study recently showed increased hippocampal size in expert musicians (14). More convincing evidence for hippocampal change is provided by a longitudinal study showing learning-associated increases in the hippocampus in a group of medical students following intensive study for examinations (8). These processes may be more efficient in the expert group of pilots than in the less expert groups of pilots. It was notable, however, that the groups of pilots did not differ on standard measures of learning and memory (such as the auditory verbal learning test, which measures declarative memory function).

Research on experts' memories for stimuli relevant to their domain (e.g., chess pieces arranged as in an actual game) has led to the concept that experts more efficiently encode domain-relevant information in the form of "chunks" (13). Aviation communications (e.g., navigational headings and altitudes) are part of a domain-specific language that has been acquired through flight training and experience. In this context, expert pilots are able to read



back and execute navigational headings and altitudes more accurately than less highly trained pilots or nonpilots (28). Advanced training in the use of navigational instruments and greater experience with ATC clearances may engender a more accurate mental representation of an assigned heading and altitude, possibly as both auditory and visuospatial representations in memory. Results from our experiment showed that the high expertise group of pilots consistently performed better than the lower expertise groups on the landing approach component of the flight simulator task (it was this component of the flight simulator task that accounted for the majority of the difference between the more and less highly trained pilots over the three annual tests). The landing approach always started at an altitude of 1200 ft (366 m) with a heading of 360°. Expert pilots having larger hippocampal size may retain and integrate this fact into their mental representation of the flight (24). Note that traffic avoidance showed the strongest main effect of hippocampal volume such that smaller hippocampal volume was associated with improved performance with repeated testing. This counterintuitive result may arise from the performances of the older pilots, especially the less expert older pilots who tend to have smaller hippocampal volumes. In the larger Stanford/VA Aviation Study, older age was associated with improved performance over time on the traffic avoidance subtask (26). Thus, the influence of age, both on improvements in the traffic avoidance task and on decline in hippocampal volume, may have led to a nonsignificant association (i.e., main effect) of hippocampal volume on overall simulator performance.

Although expert pilots' longitudinal performance on the flight simulator task partly depends on declarative memory, piloting an aircraft is a complex task of flexible attention allocation (4) involving coordination of both physical and mental abilities. Advanced flight training and experience help a pilot to allocate attention more efficiently because the pilot knows "what to look for, when to look for it, and what response to make" (9). Thus, flight simulator performance also likely depends on nondeclarative forms of memory which include motor skills and habit learning, simple forms of conditioning, and the phenomenon of priming, which are all supported by brain systems outside the hippocampus (24). Compared to declarative memory, nondeclarative memory is relatively inflexible, operates outside of awareness, and is most accessible when the task is structured just as during the original learning conditions (23). Many of these properties may be related to skills involved in flying. For example, expert motor skills are obviously important and are typically associated with brain regions, including the striatum, cerebellum, basal ganglia, prefrontal cortex, and parietal lobes (7). One model based on functional brain imaging has suggested that when a motor skill has been learned well, the neural representation maintaining this motor behavior is distributed over a neural network (7).

In order to clarify the nature of the relationship between hippocampal volume and the acquisition and maintenance of expert performance in a skill domain, studies need to measure hippocampal volumes at multiple time points. Woollett and Maguire (27) recently reported longitudinal results from a study in London taxi drivers training to become licensed. In the taxi drivers who qualified, the acquisition of an internal spatial representation of London was associated with a selective increase in gray matter volume in their posterior hippocampi.

Because of the key role of deliberate practice, future research on brain mechanisms involved in acquisition and maintenance of experts' superior performance should measure specific types and hours of deliberate practice to address how deliberate practice relates to brain structure and function. Future studies, however, should also emphasize nondeclarative memory to understand its contribution to driving and flying. Broadly, one can look to the adult human hippocampus as a target brain area where cognitive training and continuous and deliberate training can help restore, maintain, and improve performance on many types of complex everyday tasks that are initially dependent on proper hippocampal functioning. For instance, using measurement of hippocampal function may help develop individualized training programs and tools that can help rehabilitate performance impairments due to disease or injury (such as post traumatic stress disorder/traumatic brain injury). A similar approach can identify people at risk for accidents.

The present study had a few limitations. First, the relatively small sample size consisted mostly of men. The number of women should be increased in future studies. Second, effects related to aging complicated the interpretation of the results. Any expertise/experience-related associations in their study are superimposed on the large individual differences in age-related cognitive and brain processes that are occurring in their sample. As hippocampal volume varies with age, some of the variance in hippocampal volumes was also related to age. Also, our structural scans were not taken at the same point in the study for all subjects. Ideally these scans would be taken immediately prior to baseline in a cohort of pilots all near the same age. Despite the less than ideal timing of the hippocampal measurements, hippocampal volume and APOE genotype measures helped to explain individual differences in longitudinal flight simulator performance beyond that explained by pilot age and expertise. Previously, 11% of the variance was explained on the basis of pilot demographics (26). Adjusting for baseline performance explained an additional 7% of the variance. Inclusion of the two neurobiological measures explained an additional 9% variance (a total of 27% was the total percentage of variance explained).

In conclusion, the present study adds to the existing literature on expertise and the biological correlates of acquired knowledge. Expertise moderated the influence of hippocampal volume on longitudinal flight simulator performance such that having a larger hippocampus was progressively associated with better maintenance of performance as the level of expertise increased. Expert pilots with larger hippocampi were able to perform relatively better across annual testing. This finding resonates with previous reports of hippocampal volume differences in other fields of expertise, including spatial navigation (21), musical ability (14), and intensive medical study (8). However, hippocampal volume in our study was not associated with better maintenance of performance over time in the lower expertise groups. This finding suggests a complex interaction between expertise, hippocampal volume, and performance. Consideration of the existing literature suggests flight simulator performance involves interplay between declarative (hippocampal) and nondeclarative (nonhippocampal) memory systems.

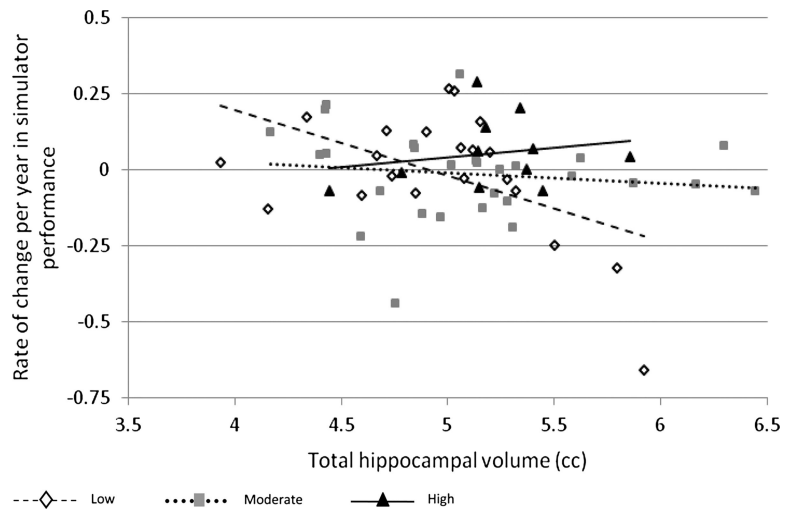
## ACKNOWLEDGMENTS

This work was supported by the Sierra-Pacific Mental Illness Research, Education, and Clinical Center (MIRECC), the War Related Illness and Injury Study Center (WRIISC), and the Medical Research Service of the Department of Veterans Affairs; and by the National Institutes of Health [grant numbers R01 AG021632 (with Diversity Supplement), P41 RR023953, and R37 AG12713]. Dr. Scanlon was supported by the Office of Academic Affiliations, Advanced Fellowship Program in Mental Illness Research and Treatment, Department of Veteran Affairs. We thank Greer Murphy, Ph.D., for providing the genetic analysis, Booil Jo, Ph.D., for statistical consultation, and Daniel Heraldez, Katy Castile, and Gordon Reade for testing subjects. We also express appreciation to the aviator study subjects for their time and interest in pursuit of answering intellectual questions. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

## REFERENCES

1. Adamson MM, Hutchinson JB, Shelton AL, Wagner AD, Taylor JL. Reduced hippocampal activity during encoding in cognitively normal adults carrying the APOE e4 allele. *Neuropsychologia* 2011; 49: 2448–55. [PubMed: 21549723]
2. Adamson MM, Landy KM, Duong S, Fox-Bosetti S, Ashford JW, et al. Apolipoprotein E epsilon4 influences on episodic recall and brain structures in aging pilots. *Neurobiol Aging* 2010; 31: 1059–63. [PubMed: 18760504]
3. Adamson MM, Samarina V, Xiangyan X, Huynh V, Kennedy Q, et al. The impact of brain size on pilot performance varies with aviation training and years of education. *J Int Neuropsychol Soc* 2010; 16: 412–23. [PubMed: 20193103]
4. 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: 569–79. [PubMed: 9215461]
5. Boyke J, Driemeyer J, Gaser C, Buchel C, May A. Training-induced brain structure changes in the elderly. *J Neurosci* 2008; 28: 7031–5. [PubMed: 18614670]
6. Cabeza R, Ciaramelli E, Olson IR, Moscovitch M. The parietal cortex and episodic memory: an attentional account. *Nat Rev Neurosci* 2008; 9: 613–25. [PubMed: 18641668]
7. Doyon J, Benali H. Reorganization and plasticity in the adult brain during learning of motor skills. *Curr Opin Neurobiol* 2005; 15: 161–7. [PubMed: 15831397]
8. Draganski B, Gaser C, Kempermann G, Kuhn HG, Winkler J, et al. Temporal and spatial dynamics of brain structure changes during extensive learning. *J Neurosci* 2006; 26: 6314–7. [PubMed: 16763039]
9. Durso FT, Dattel AR. Expertise and transportation In: Ericsson KA, Charness N, Feltovich P, Hoffman R, eds. *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press; 2006:355–71.
10. Ericsson KA, Krampe RT, Tesch-Römer C. The role of deliberate practice in the acquisition of expert performance. *Psychol Rev* 1993; 100:363–406.
11. Instrument Flying Handbook FAA. Washington, DC: U.S. Department of Transportation; 2008:11.
12. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, et al. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 2002; 33: 341–55. [PubMed: 11832223]
13. Gobet F, Charness N. Chess and games *Cambridge handbook on expertise and expert performance*. Cambridge, MA: Cambridge University Press; 2006: 523–38.
14. Groussard M, La Joie R, Rauchs G, Landeau B, Chételat G, et al. When music and long-term memory interact: effects of musical expertise on functional and structural plasticity in the hippocampus. *PLoS ONE* 2010; 5: e13225. [PubMed: 20957158]
15. Hales JB, Brewer JB. Activity in the hippocampus and neocortical working memory regions predicts successful associative memory for temporally discontinuous events. *Neuropsychologia* 2010; 48: 3351–9. [PubMed: 20667491]
16. Herdener M, Esposito F, di Salle F, Boller C, Hilti CC, et al. Musical training induces functional plasticity in human hippocampus. *J Neurosci* 2010; 30: 1377–84. [PubMed: 20107063]

17. Hill NM, Schneider W. Brain changes in the development of expertise: neuroanatomical and neurophysiological evidence about skill-based adaptations In: Ericsson KA, Charness N, Feltovich PJ, Hoffman RR, eds. *The Cambridge handbook of expertise and expert performance*. New York: Cambridge University Press; 2006:653–82.
18. Kirk U, Skov M, Christensen MS, Nygaard N. Brain correlates of aesthetic expertise: a parametric fMRI study. *Brain Cogn* 2009; 69: 306–15. [PubMed: 18783864]
19. Klucharev V, Smidts A, Fernandez G. Brain mechanisms of persuasion: how ‘expert power’ modulates memory and attitudes. *Soc Cogn Affect Neurosci* 2008; 3: 353–66. [PubMed: 19015077]
20. Lycke C, Specht K, Ersland L, Hugdahl K. An fMRI study of phonological and spatial working memory using identical stimuli. *Scand J Psychol* 2008; 49: 393–401. [PubMed: 18705673]
21. Maguire EA, Woollett K, Spiers HJ. London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. *Hippocampus* 2006; 16: 1091–101. [PubMed: 17024677]
22. Miyashita Y Cognitive memory: cellular and network machineries and their top-down control. *Science* 2004; 306: 435–40. [PubMed: 15486288]
23. Squire LR, Clark RE, Bayley PJ. Medial temporal lobe memory function and memory In: Gazzaniga MS, ed. *The cognitive neurosciences*, 3rd ed. Cambridge, MA: The MIT Press; 2004: 691–708.
24. Squire LR, Stark CE, Clark RE. The medial temporal lobe. *Annu Rev Neurosci* 2004; 27: 279–306. [PubMed: 15217334]
25. Taylor JL, Kennedy Q, Adamson MM, Lazzeroni LC, Noda A, et al. Influences of APOE epsilon4 and expertise on performance of older pilots. *Psychol Aging* 2011; 26: 480–7. [PubMed: 21668123]
26. Taylor JL, Kennedy Q, Noda A, Yesavage JA. Pilot age and expertise predict flight simulator performance: a 3-year longitudinal study. *Neurology* 2007; 68: 648–54. [PubMed: 17325270]
27. Woollett K, Maguire EA. Acquiring “the knowledge” of London’s layout drives structural brain changes. *Curr Biol* 2011; 21: 2109–14. [PubMed: 22169537]
28. Yesavage JA, Jo B, Adamson MM, Kennedy Q, Noda A, et al. Initial cognitive performance predicts longitudinal aviator performance. *J Gerontol B Psychol Sci Soc Sci* 2011; 66: 444–53. [PubMed: 21586627]
29. Murphy GM, Taylor J, Kraemer HC, Yesavage J, Tinklenberg JR. No association between apolipoprotein E epsilon 4 allele and rate of decline in Alzheimer’s disease. *Am J Psychiatry* 1997; 154: 603–8. [PubMed: 9137113]



**Fig. 1.** Relationship between hippocampal volume and longitudinal flight simulator performance. Trend lines are shown for each expertise group. Y-axis: annualized rate of change in flight simulator performance from baseline to 2 yr later (slope score residualized on baseline flight simulator score). X-axis: hippocampal volume, adjusted for head size (TIV). The sample mean, 5.058 cc, has been added to the residualized hippocampal volume.

**TABLE I.**  
 DEMOGRAPHIC CHARACTERISTICS (MEAN  $\pm$  SD) OF THE 60 SUBJECTS BY LEVEL OF EXPERTISE<sup>1</sup>.

	Least (VFR)	Moderate (IFR)	Most (CFII/ATP)
	N = 21	N = 28	N = 11
Age, yr, mean $\pm$ SD.	56.9 $\pm$ 7.5 (age range = 46–69)	59.0 $\pm$ 5.4 (age range = 49–69)	53.4 $\pm$ 5.5 (age range = 46–65)
Education, yr, mean $\pm$ SD.	16.5 $\pm$ 1.9	17.3 $\pm$ 1.4	18.1 $\pm$ 2.1
Number Men	16 (76%)	22 (79%)	11 (100%)
Number (%) APOE $\epsilon$ 4 carriers <sup>2</sup>	11 (52%)	10 (36%)	5 (45%)
Number Employed/Unemployed/Retired	17/2/2	17/2/9	10/1/0
Total hours of flight experience, mean $\pm$ SD.	719 $\pm$ 573	1631 $\pm$ 1290	3228 $\pm$ 1616
Total flight hours in past month, mean $\pm$ SD <sup>3</sup>	4.2 $\pm$ 4.3	12.6 $\pm$ 15.0	22.6 $\pm$ 25.6
Flight Simulator Performance at baseline (z-score composite) <sup>*</sup>	-0.09 $\pm$ 0.68	-0.01 $\pm$ 0.47	0.39 $\pm$ 0.34
Total Intracranial Volume (TIV; cc)	1685.3 $\pm$ 164.3	1694.5 $\pm$ 181.2	1806.1 $\pm$ 175.4
Hippocampal Volume (TIV; cc) <sup>4</sup>	4.99 $\pm$ 0.59	5.11 $\pm$ 0.56	5.11 $\pm$ 0.34
Rey AVLT (z-score composite)	0.03 $\pm$ 0.92	-0.12 $\pm$ 0.97	0.39 $\pm$ 0.82

<sup>\*</sup>  $P < 0.05$ .

<sup>1</sup> Expertise was defined FAA pilot proficiency ratings collected at entry in the study.

<sup>2</sup> APOE genotyping is based on genomic DNA extracted from frozen blood/buccal mucosa/saliva samples based on Murphy et al. (29). All subjects agreed to have the results of APOE genotyping withheld from them (APOE  $\epsilon$ 4/4 = 2; APOE  $\epsilon$ 3/4 = 24; APOE  $\epsilon$ 3/3 = 34).

<sup>3</sup> Flight hours past month: number of hours flown (nonsimulator) in the month immediately preceding the baseline visit.

<sup>4</sup> Hippocampal volumes (in cc units) were residualized on total intracranial volume (TIV) to adjust for head size.

TABLE II.

RELATIONSHIP OF HIPPOCAMPAL VOLUME, AGE, AVIATION EXPERTISE, APOE ε4 CARRIER STATUS AND BASELINE FLIGHT SIMULATOR PERFORMANCE WITH LONGITUDINAL FLIGHT SIMULATOR PERFORMANCE IN 60 AVIATORS: SUMMARY OF MULTIPLE REGRESSION MAIN “NON-ADJUSTED” MODEL RESULT (PARAMETER ESTIMATES ± SE AND OVERALL R<sup>2</sup>).

Main Model: Expertise Rating at Baseline										
Flight Summary Score <sup>1</sup>	Intercept (β0)	Hippocampal Volume <sup>2</sup> (β1)	Age (β2)	Expertise <sup>3</sup> (β3)	APOE ε4 Status (β4)	Baseline Flight Summary Score <sup>1</sup> (β5)	Expertise × Hippocampal Volume (β6)	Expertise × ε4 Status (β7)	Overall F(7, 52)	Overall R <sup>2</sup>
	-0.068 ± 0.022**	-0.055 ± 0.046	-0.006 ± 0.004	0.051 ± 0.031	-0.032 ± 0.044	-0.099 ± 0.045	0.182 ± 0.072	0.116 ± 0.059	2.70*	0.267

\* P < 0.05

\*\* P < 0.01.

<sup>1</sup> Longitudinal overall flight simulator performance is indexed by the slope of the “Flight summary score,” which is a composite z-score of the flight tasks [communications, avoidance, emergency, and approach; see Taylor (26) for details]. Baseline Flight Summary Score is calculated as the composite z-score of the four flight tasks [communications, avoidance, emergency, and approach; see Taylor (26) for details] on the first visit.

<sup>2</sup> Hippocampal volumes (in cc units) were residualized on total intracranial volume (TIV) to adjust for head size.

<sup>3</sup> Expertise is measured by FAA pilot proficiency ratings, which provide an ordinal measure of lifetime aviation training and experience. This measure was collected at baseline.

The parameters estimated by the general linear model for predicted rate of change per year in a flight simulator: summary z-s score = -0.068 - 0.055\*(hippocampal volume) + 0.051\*(expertise rating) + 0.182\*(expertise rating)\*(hippocampal volume) + 0.116\*(expertise rating)\*(APOE ε4 status) - 0.006\*(average age for each rating group) - 0.99\*(baseline flight simulator summary z-score). Expertise rating of -1, 0, and 1 represents least, moderate, and most expertise, respectively.



**TABLE III.**  
**SCORES AMONG STUDY VARIABLES, INCLUDING DEMOGRAPHICS, APOE ε4 STATUS, FLIGHT SIMULATOR PERFORMANCE, COMPOSITE SCORES AND LOBAR VOLUMES\* .**

Age	Years of Education	APOE ε4 Status	Flight Hours Past Month	Accumulated Flight Hours	Aviation Expertise	Flight Performance Baseline	Flight Performance Slope (2 yr)	Rey AVLT Composite Baseline	Hippocampus	Parietal Lobe Volume	Temporal Lobe Volume	Occipital Lobe Volume	Basal Ganglia Volume	Cerebellar Volume	Prefrontal Cortex Volume	Whole Brain Volume	TIV
0.07		-0.07	0.01	-0.15	-0.13	-0.41	-0.01	-0.36	-0.32	-0.35	-0.27	-0.15	-0.36	-0.05	-0.42	-0.39	0.16
0.07		0.24	0.06	0.01	<b>0.32</b>	0.05	-0.06	0.13	0.01	-0.03	-0.03	-0.14	-0.05	-0.02	-0.03	-0.09	0.06
-0.07	0.24		0.03	-0.14	-0.08	0.03	-0.10	-0.20	0.00	0.05	0.19	0.01	0.12	0.03	0.08	0.13	0.07
0.01	0.06	0.03		<b>0.82</b>	<b>0.40</b>	0.14	0.18	0.02	0.01	-0.06	-0.01	0.05	<b>0.28</b>	-0.04	-0.11	-0.06	<b>0.38</b>
-0.15	0.01	-0.14	<b>0.82</b>		<b>0.37</b>	0.23	0.15	0.13	0.07	0.03	0.01	0.07	<b>0.26</b>	0.02	-0.07	0.05	<b>0.27</b>
-0.13	<b>0.32</b>	-0.08	<b>0.40</b>	<b>0.37</b>		<b>0.28</b>	0.08	0.10	0.17	-0.13	-0.06	-0.08	0.24	-0.04	0.02	-0.09	0.21
-0.41	0.05	0.03	0.14	0.23	<b>0.28</b>		-0.16	0.18	0.07	<b>0.26</b>	0.20	<b>0.31</b>	<b>0.32</b>	0.22	<b>0.32</b>	<b>0.34</b>	0.06
-0.01	-0.06	-0.10	0.18	0.15	0.08	-0.16		-0.12	-0.17	0.10	0.21	0.11	0.01	-0.13	0.17	0.06	0.12
-0.36	0.13	-0.20	0.02	0.13	0.10	0.18	-0.12		<b>0.30</b>	0.19	0.13	0.12	<b>0.30</b>	0.07	0.18	0.14	<b>-0.32</b>
-0.32	0.01	0.00	0.01	0.07	0.17	0.07	-0.17	<b>0.30</b>		0.21	0.19	0.19	<b>0.28</b>	0.16	0.13	0.26	0
-0.35	-0.03	0.05	-0.06	0.03	-0.13	<b>0.26</b>	0.10	0.19	0.21		<b>0.74</b>	<b>0.53</b>	<b>0.47</b>	0.21	<b>0.70</b>	<b>0.81</b>	0
-0.27	-0.03	0.19	-0.01	0.01	-0.06	0.20	0.21	0.13	0.19	<b>0.74</b>		<b>0.46</b>	<b>0.41</b>	0.17	<b>0.67</b>	<b>0.73</b>	0
-0.15	-0.14	0.01	0.05	0.07	-0.08	<b>0.31</b>	0.11	0.12	0.19	<b>0.53</b>	<b>0.46</b>		<b>0.33</b>	-0.01	<b>0.46</b>	<b>0.50</b>	0

*Aviat Space Environ Med.* Author manuscript; available in PMC 2020 May 05.

Age	Years of Education	APOE ε4 Status	Flight Hours Past Month	Accumulated Flight Hours	Aviation Expertise	Flight Performance Baseline	Flight Performance Slope (2 yr)	Rev AVLT Composite Baseline	Hippocampus	Parietal Lobe Volume	Temporal Lobe Volume	Occipital Lobe Volume	Basal Ganglia Volume	Cerebellar Volume	Prefrontal Cortex Volume	Whole Brain Volume	TIV
-0.36	-0.05	0.12	<b>0.28</b>	<b>0.26</b>	0.24	<b>0.32</b>	0.01	<b>0.30</b>	<b>0.28</b>	<b>0.47</b>	<b>0.41</b>	<b>0.33</b>	0.13	0.13	<b>0.48</b>	<b>0.54</b>	0
-0.05	-0.02	0.03	-0.04	0.02	-0.04	0.22	-0.13	0.07	0.16	0.21	0.17	-0.01	0.13	0.07	0.07	0.20	0
<b>-0.42</b>	-0.03	0.08	-0.11	-0.07	0.02	<b>0.32</b>	0.17	0.18	0.13	<b>0.70</b>	<b>0.67</b>	<b>0.46</b>	<b>0.48</b>	0.07	<b>0.81</b>	<b>0.81</b>	0
<b>-0.39</b>	-0.09	0.13	-0.06	0.05	-0.09	<b>0.34</b>	0.06	0.14	0.26	<b>0.81</b>	<b>0.73</b>	<b>0.50</b>	<b>0.54</b>	0.20	<b>0.81</b>	<b>0.81</b>	0
0.16	0.06	0.07	<b>0.38</b>	<b>0.27</b>	0.21	0.06	0.12	<b>-0.32</b>	0	0	0	0	0	0	0	0	0

on coefficients in bold were significant ( $P < 0.05$ ).

at the time of the baseline flight simulator performance.

on are truncated to 20 years.

mping is based on genomic DNA extracted from frozen blood/buccal mucosa/saliva samples.

month: number of hours flown (nonsimulator) in the month immediately preceding the baseline visit.

ight Hours is the average yearly increase in flight hours during the course of follow-up. At each annual follow-up, the number of flight hours accumulated to date were recorded, based on verification of the pilot's log book whenever possible.

se: measured by FAA proficiency ratings at entry into the study which provide an ordinal measure of lifetime aviation training and experience.

performance: a composite z-score of the four flight tasks; (see Taylor 2007 for details).

performance: linear slopes of the flight performance from baseline to year 2 visit; [see Taylor (26) for details].

posite score is the average of the immediate and delayed recall z-scores [see Adamson et al. (1) for details].

olumes were residualized on total intracranial volume (TIV) to adjust for head size.

retal, temporal, occipital, cerebellum, prefrontal cortex, and whole brain were residualized on TIV to adjust for head size.

cranial volume (cc).