ORIGINAL RESEARCH

Relationship of CogScreen-AE to Flight Simulator Performance and Pilot Age

JOY L. TAYLOR, PH.D., RUTH O'HARA, PH.D., MARTIN S. MUMENTHALER, PH.D., AND JEROME A. YESAVAGE, M.D.

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; 71:373–80.**

Objectives: We report on the relationship between CogScreen-Aeromedical Edition (AE) factor scores and flight simulator performance in aircraft pilots aged 50–69. ***Methods:*** Some 100 licensed, civilian aviators (average age 58 ± 5.3 yr) performed aviation tasks in a Frasca model 141 flight simulator and the CogScreen-AE battery. The aviation performance indices were: a) staying on course; b) dialing in communication frequencies; c) avoiding conflicting traffic; d) monitoring cockpit instruments; e) executing the approach; and f) a summary score, which was the mean of these scores. The CogScreen predictors were based on a factor structure reported by Kay (11), which comprised 28 CogScreen scores. Through principal components analysis of Kay's nine factors, we reduced the number of predictors to five composite CogScreen scores: Speed/Working Memory (WM), Visual Associative Memory, Motor Coordination, Tracking, and Attribute Identification. ***Results:*** Speed/WM scores had the highest correlation with the flight summary score, Spearman rrho = 0.57. A stepwise-forward multiple regression analysis indicated that four CogScreen variables could explain 45% of the variance in flight summary scores. Significant predictors, in order of entry, were: Speed/WM, Visual Associative Memory, Motor Coordination, and Tracking ($p < 0.05$). Pilot age was found to significantly improve prediction beyond that which could be predicted by the four cognitive variables. In addition, there was some evidence for specific ability relationships between certain flight component scores and CogScreen scores, such as approach performance and tracking errors. ***Conclusions:*** These data support the validity of CogScreen-AE as a cognitive battery that taps skills relevant to piloting.

Keywords: aerospace medicine, cognitive assessment, aging relationship, CogScreen-AE, flight simulator performance, pilot age.

IN RECOGNIZING A NEED for cognitive evaluations of licensed pilots (1), the Federal Aviation Administration (FAA) sponsored the development of CogScreen-Aeromedical Edition (AE) (11). The CogScreen battery was designed to measure the underlying perceptual, cognitive, and information processing abilities associated with flying, and to provide a sensitive and specific instrument for use in the medical recertification evaluation of pilots with known or suspected neurological and/or neuropsychiatric conditions. Preliminary data suggests that CogScreen-AE performs comparatively well in discriminating between neurologically impaired and cognitively intact individuals (11). To establish the validity of CogScreen as an occupationally relevant assessment instrument, it is crucial to demonstrate a relationship between CogScreen scores and flight performance. This paper documents such a relationship, extending the results of three recent studies addressing this issue (9,10,26). Below we describe our analytic approach and major hypotheses.

CogScreen and Flight Performance

A problem with CogScreen-AE, especially for research, is how to deal with the myriad of scores generated by the battery. There are potentially 65. If all of the individual CogScreen variables are used in hypothesis testing or in prediction work, then the speed scores (correct reaction time) and thruput scores (number of correct reactions per unit of time) arising from a given CogScreen test are very likely to be collinear simply because of the arithmetic relation between speed and thruput (this especially applies when accuracy rates approach 100%). Second, some of the individual CogScreen tests may measure essentially the same cognitive ability, and may for this reason be collinear. Therefore, in our examination of the relationship between CogScreen and flight simulator performance, we felt it would be advantageous to combine the individual CogScreen variables into a smaller set of relatively nonredundant variables. As a starting point, we used a preliminary factor structure reported by Kay (11), which is based on 28 CogScreen variables. The factor names, or abilities thought to be measured, are: a) Visual Scanning/Sequencing; b) Attribute Identification; c) Visual Perceptual/Spatial Processing; d) Motor Coordination; e) Choice Visual Reaction Time; f) Visual Associative Memory; g) Tracking; h) Working Memory; and i) Numerical Operations.

The breadth of 13 subtests of CogScreen-AE allows an examination of CogScreen-flight performance relationships within specific aviation performance domains.

From the Department of Veterans Affairs Health Care System, Palo Alto, CA (J. L. Taylor, J. A. Yesavage); and Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA (R. O'Hara, M. S. Mumenthaler, J. A. Yesavage).

This manuscript was received for review in February 1999. It was revised and accepted for publication in August 1999.

Address reprint requests to: Joy L. Taylor, Ph.D., who is Assistant Director of the Dementia Core of the Sierra-Pacific Mental Illness Research, Education, and Clinical Center, VA Palo Alto Health Care System (151Y), 3801 Miranda Avenue, Palo Alto, CA 94304-1207; joyt@stanford.edu

For instance, CogScreen-AE has a psychomotor tracking measure that may correlate with approach and landing performance. There is a spatial processing test, Manikin, which may correlate with the ability to react quickly and appropriately to conflicting traffic. Finally, the Visual Scanning/Sequencing factor may predict performance in scanning cockpit instruments.

To investigate the relationship between CogScreen-AE and flight-simulator performance, 100 licensed civilian pilots, aged 50–69 yr, were administered the CogScreen-AE battery along with aviation tasks performed in a small-aircraft flight simulator. The flight simulator scenario presents a variety of routine and non-routine aviation tasks including take-off and landing maneuvers, scanning cockpit instruments, avoiding conflicting air traffic, responding to air-traffic controller (ATC) communications, and navigating ATC-assigned courses. Performance on several of these tasks has been found to be sensitive to the influence of age (16,23,28) and psychoactive drugs, including marijuana, ethanol, and nicotine (13,15,17,22,27). Furthermore, scores on the ATC Communications task have been found to correlate with WAIS-R Backward Digit Span (23).

CogScreen and Pilot Age

Because there are well-documented age-related differences in performance of cognitive tests and in view of the controversy surrounding the Age-60 Rule (FARS 121.383c), we also looked at CogScreen factor scores in relation to age: Where and how much do younger and older pilots differ in CogScreen performance? Is flight simulator performance predicted better by cognitive ability or by pilot age? We expected that one or more CogScreen variables would provide a speed-of-processing measure, that speed would significantly correlate with overall flight performance and with age, and that speed would account for as much or more variance in flight performance than age. While speed of processing as a predictor of pilot performance has received less attention in aviation psychology than time-sharing (2,18,24) and specialized ability approaches (6,7), some cognitive aging theorists suggest that age-related decline in cognition may be mediated by age-related decline in speed of processing (14,20). Hence, if aviators of various ages differ in their speed of processing, then speed should predict cognitively demanding aviation performance.

METHODS

Subjects

Some 100 pilots holding current FAA medical certificates were recruited into an ongoing longitudinal study of civilian pilots between 50–69 yr of age. Pilots were required to be actively flying and to have at least 300 h of total flight experience, but no more than 15,000 h, at entry. To obtain a relatively homogeneous sample in terms of past and current flight experience, pilots who had ever flown for major air carriers were excluded from participating. At entry, 23% were private-licensed pilots rated for visual flight conditions, 66% were non-air-transport instrument-rated pilots, and the remaining 11% held air-transport ratings. The majority were medically certified as Airman Class III (52%), 40% were Class II, and 8% were certified as Class I. The sample consists of 82 men and 18 women, with an average of 16.7 yr of education (SD = 2.1). Some 95% are Caucasian (non-Hispanic), 3% African-American, and 2% reported other ethnic/racial backgrounds. All subjects gave written informed consent to participate and could withdraw at any time. The protocol was approved by the Human Subject Committee of Stanford University and has been performed in accordance with the ethical standards established by the 1964 Declaration of Helsinki.

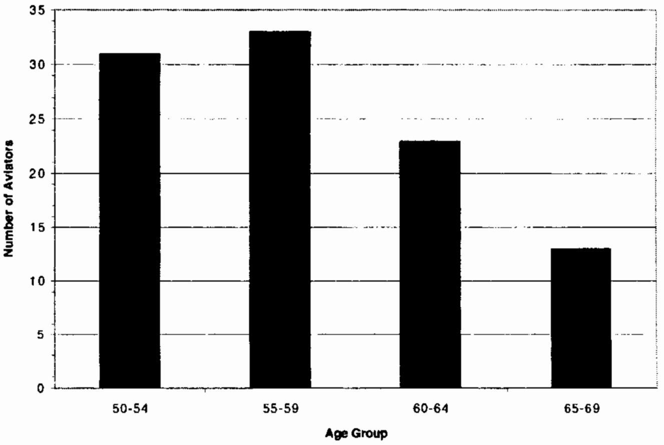

Fig. 1. Age distribution of the sample (Mean age = 57.9 yr; n = 100).

The data reported here were collected at study entry when the average age was 57.9 yr (SD = 5.3; range = 50–69) and the average hours of flight logged was 2464 h (SD = 2498; range = 310–11,800 h). **Fig. 1** presents the age distribution. Pilot age was not significantly correlated with total hours of flight experience or with years of education (Spearman r_{rho} = 0.13 and −0.11, respectively, $p > 0.15$), but subjects with more education tended to have fewer flight hours, $r_{rho} = -0.21$, $p < 0.05$.

Equipment and Measures

*Flight simulator:*A Frasca 141 flight simulator (Urbana, IL) was linked to an IRIS 4D computer (Silicon Graphics, Mountain View, CA) that generated realistic "through-the-window" graphics of the environment and collected data concerning the aircraft's flight conditions. The equipment simulated flying a small single-engine aircraft with fixed landing gear and fixed propeller above flat terrain with surrounding mountains and clear skies. A speaker system was installed in the cockpit and connected to a tape recorder, through which the pilot received Air Traffic Control (ATC) messages that were in accordance with FAA standards (FAA Order 7110.650). Each ATC script contained a take-off clearance, 16 critical en route messages and instructions for approach and landing.

Each flight lasted 75 min and consisted of a standard scenario with 19 legs around the airport, including leg 1: take-off, legs 2–17: en route flying, leg 18: approach, leg 19: landing. During en route flying, pilots were given a new ATC command every 3 min with new heading, altitude, radio frequency, and in 50% of the legs, a new transponder code. In order to increase the pilots' workload on legs 2–17, we confronted them with three different emergency situations: carburetor icing, drop of engine oil pressure, and suddenly approaching air traffic, which occurred randomly (total of 19 occurrences in 48 min). Pilots were trained to immediately report the engine emergency and to avoid the oncoming traffic by veering quickly yet safely in a direction diagonal to the traffic. Throughout the flight pilots flew in severe turbulence; there was also a 15-kn crosswind at a 90° angle to the airstrip during approach and landing. To increase the difficulty of an aligned take-off and landing, our scenario did not provide a marked centerline and touchdown point on the runway.

CogScreen-AE: The entire battery was administered. Further details can be found in the manual (11) and in prior articles of this journal (4,12).

Procedures

To minimize practice effects and achieve relatively stable and reliable flight simulator performance, each subject participated in six simulator training sessions prior to the test day. In these training sessions aviators performed the same flight tasks to be performed on the test day, but a new set of ATC communications was presented each time. On the test day, subjects flew one morning and one afternoon flight. After each flight, they spent 40–60 min taking cognitive ability tests. By random assignment, subjects performed the CogScreen battery after either the morning or the afternoon flight. The test day lasted approximately 6 h, including a 45- to 60-min lunch break.

TABLE I. MEAN (±SD) SCORES ON THE 28 COGSCREEN-AE VARIABLES OF THE KAY (1995) FACTOR ANALYSIS.

Measure	N	Mean	SD
Pathfinder Letter Speed	100	0.8	0.2
Pathfinder Number Speed	100	1.0	0.3
Pathfinder Combined Speed	100	1.4	0.5
Shifting Attention Arrow Direction Thruput	86	96.4	21.1
Shifting Attention Rule Shift Comp	84	6.0	3.0
Shifting Attention Discovery Accuracy	84	61.2	14.7
Shifting Attention Failures to Maintain Set	84	2.3	1.8
Shifting Attention Perseverative Errors	84	2.5	3.9
Shifting Attention Sequence Thruput	98	22.5	6.0
Visual Sequence Comparison Thruput	100	23.0	5.6
Matching-to-Sample Thruput	100	41.3	7.4
Symbol Digit Coding Thruput	100	26.5	6.4
Manikin Thruput	100	32.0	8.8
Shifting Attention Instruction Thruput	86	67.6	13.5
Pathfinder Number Coordination	100	0.9	0.3
Pathfinder Letter Coordination	100	1.1	0.3
Pathfinder Combined Coordination	100	5.9	48.7
Divided Attention Indicator Alone Speed	100	0.4	0.1
Divided Attention Indicator Dual Speed	100	0.9	0.4
Shifting Attention Arrow Color Thruput	86	86.2	17.1
Symbol Digit Coding Immed. Recall Accuracy	100	79.2	25.0
Symbol Digit Coding Delayed Recall Accuracy	100	69.5	27.8
Dual Task Tracking Alone Error	99	23.6	15.7
Dual Task Tracking Dual Error	98	71.7	23.1
Dual Task Prev Num Alone Thruput	97	102.8	63.0
Dual Task Prev Num Dual Thruput	93	74.9	44.1
Math Thruput	100	2.1	0.8
Backward Digit Span Accuracy	100	83.6	15.9

Note: The smaller numbers for the Shifting Attention scores are because of a technical problem during test administration that affected the earlier enrollees in this study.

Data Reduction

Flight simulator: The flight performance indices in this study are based on the average of the morning and afternoon flights, which tends to increase the reliability of the variables analyzed. Six scores, reflecting performance after takeoff and before landing, were analyzed. These performance indices related to: a) staying on course; b) dialing in communication frequencies; c) avoiding conflicting traffic; d) monitoring cockpit instruments; e) executing the approach; and f) a summary score, which was the mean of scores 1–5. Retest reliability coefficients of these measures were in the 0.55 to 0.80 range. The take-off and landing scores were not analyzed because the reliability coefficients were less than 0.50. While the reason(s) for low reliability of the take-off and landing scores are not known, possible reasons include restricted range of performance and the visually degraded display of the runway. If the loss of horizontal and vertical distance information was substantial, pilots might have "guessed" the positions of the centerline and ideal touchdown point.

The flight scores in this work and our previous work are z-scores since the raw data have different units of measurement (e.g., altitude in feet, heading in degrees, airspeed in knots, reaction time in seconds). Initially, the scoring system of the flight simulator-computer unit produces 23 raw scores derived from response latencies, errors in communication frequencies, or deviations from ideal or assigned positions. From the raw scores we compute z-scores, using the sample mean and SD for each variable. The 23 standardized variables are then aggregated into flight component scores (course, communication frequencies, traffic avoidance, cockpit monitoring and approach) plus the flight summary score. More detailed descriptions of the flight scenario and scoring are available in Taylor et al. (22).

CogScreen: Kay (11), using the normative data set of some 600 commercial airline pilots, has proposed preliminary factor structures for the CogScreen measures, which comprise a total of 65 speed, accuracy, thruput, and process scores. For the 19 speed scores, Kay identified 5 factors that accounted for about 55% of the variance. Kay proposed a 9-factor structure, accounting for about 67% of the variance, for a larger set of 28 selected speed, accuracy, thruput, and process variables (named in **Table I**). We used the 9-factor structure in the present study such that we selected the same 28 variables for analysis, standardized them (with respect to the mean and SD of this sample), and then clustered them on the basis of Kay's factor structure, as summarized in **Table II**.

RESULTS

CogScreen-AE Factor Scores: Intercorrelations and Correlations with Pilot Age

Table I shows the mean raw scores (±SD) for the 28 individual CogScreen variables used in computing 9

TABLE II. AGE-PARTIALLED INTERCORRELATIONS AMONG 9 COGSCREEN-AE FACTORS PROPOSED BY KAY (1995).

Factor Name	Test Scores Included	Factor	1	2	3	4	5	6	7	8	9
Visual Scanning/ Sequencing	Pathfinder Letter, Number, and Combined speed; Shifting Attention Arrow Direction and Color thruput	1	*−0.25*	0.27	0.54	0.09	−0.57	0.25	−0.08	0.38	0.41
Attribute Identification	Shifting Attention Discovery: rule shifts completed, accuracy, failures to maintain set, perseverative errors	2		*−0.33*	0.26	0.11	−0.27	0.17	−0.28	0.43	0.19
Visual Perceptual/ Spatial Processing	Divided Attention Sequence Comparison, Visual Sequence Comparison, Matching to Sample, Symbol-Digit Coding, Manikin, and Shifting Attention Instruction thruput	3			*−0.38*	−0.11	−0.65	0.24	−0.22	0.55	0.53
Motor Coordination	Pathfinder Letter and Number coordination	4				*−0.05*	−0.05	0.07	−0.00	−0.05	0.03
Choice Visual Reaction Time	Divided Attention Indicator Alone and Dual speed; Shifting Attention Arrow Color thruput	5					*0.30*	−0.13	0.26	−0.38	−0.25
Visual Associative Memory	Symbol-Digit thruput, immediate and delayed recall	6						*−0.34*	−0.12	0.18	0.01
Tracking	Dual Task Tracking Alone and Dual error	7							*0.12*	−0.18	−0.20
Working Memory	Dual Task Previous Number Alone and Dual thruput	8								*−0.24*	0.32
Numerical Operations	Math thruput	9									*−0.21*

Notes: Age correlations are on the diagonal (italicized); Intercorrelations are above the diagonal. N = 82.
Spearman r_{rho} correlation coefficients. $|r_{rho}| \geq 0.20$ is significant at $p < 0.05$.

factor scores (11). Table II gives correlations of each factor with age (on the diagonal) and age-partialled intercorrelations of the CogScreen factor scores (above the diagonal). We note that one of the factor scores, Visual Perceptual/Spatial Processing, was significantly related to total hours of flight experience, $r_{rho} = -0.29$, $p < 0.01$, in the direction of more experienced pilots having lower thruput.

Age was significantly correlated with 7 of the 9 factor scores ($p < 0.05$). None of the age relations favored older pilots. The cognitive scores associated with age were Visual Scanning/Sequencing, Attribute Identification, Visual Perceptual/Spatial Processing, Choice Visual Reaction Time, Visual Associative Memory, Working Memory, and Numerical Operations. The factor scores not significantly related to age were Motor Coordination and Tracking. Although the statistically significant age relations were moderate (the largest being −0.38), we partialled age from the factor intercorrelations to obtain a more interpretable pattern of relations among the CogScreen-AE factors themselves. Looking at Table II, it can be seen that some of the stronger intercorrelations were among three factors relating to visual processing speed (Factors 1, 3, and 5), with $|r_{rho}|$ between 0.54 and 0.65, $p < 0.0001$. Also, Factor 3 (Visual Perceptual/Spatial Processing) was correlated with Working Memory and Numerical Operation factors, $p < 0.0001$.

Because some factors were intercorrelated, we sought to reduce the 9 factors to a set of less redundant predictors. Variable aggregation was aided by principal-components analysis (PCA) with varimax rotation. A PCA of the age-partialled correlation matrix suggested a 3-factor solution, which accounted for 61% of the variance. Loading strongly on the first factor, with loadings higher than 0.55, were the factors relating to visual processing speed (Factors 1, 3, and 5), working memory (WM; Factor 8), and math computational skill (Factor 9). We averaged the 5 factors to provide a general "Speed/WM" measure because speed of processing and working memory have been previously reported to be related (19). Furthermore, the two items used in computing the WM factor score, thruput for Previous Number (Alone condition) and thruput for Previous Number (Dual condition), were highly correlated with the speed scores for these same tasks (r_{rho} of −0.97 and −0.92, respectively). As one would expect based on the correlations of the Kay factors with age, the Speed/WM composite score was significantly correlated with age, $r_{rho} = -0.33$, $p < 0.001$, $n = 97$.

Loading strongly on the second PCA factor were Factors 2 and 7, Attribute Identification and Tracking.

TABLE III. AGE-PARTIALLED INTERCORRELATIONS AMONG 5 COGSCREEN-AE PREDICTORS.

	Attribute Identification	Tracking	Motor Coordination	Visual Associative Memory
General Speed/WM	0.36	−0.28	0.01	0.22
Attribute Identification		−0.28	0.11	0.17
Tracking			−0.00	−0.12
Motor Coordination				0.07

Notes: Except for Tracking, higher CogScreen-AE scores indicate better performance.
Spearman age-partialled correlation coefficients. $r_{rho} \geq 0.224$ is significant at $p < 0.05$.

We chose not to combine these two factors for three reasons: the bivariate correlation was modest (−0.28); tracking performance has a special importance in pilot performance assessment (2,11); and the Attribute Identification measure is based largely on performance of the Shifting Attention Discovery task, a test designed to assess executive functioning.

Loading strongly on the third factor was the Motor Coordination factor (0.86), along with a modest loading for Visual Associative Memory (0.47). Again, we opted to keep these scores separate because the intercorrelation was low (0.06) and the similarities between the two tasks are unclear. Future studies may find reason to combine Factor 4 with 6, and Factor 2 with 7. Our primary intent was to avoid strongly collinear predictors, such as was apparent with Factors 1, 3, and 5. Thus, a reduced set of five variables—General Speed/WM, Attribute Identification, Tracking, Motor Coordination, and Visual Associative Memory—was used to identify relationships between CogScreen-AE scores and flight simulator performance. Age-partialled intercorrelations among the 5 variables are presented in **Table III**.

Correlation of Reduced Set of CogScreen-AE Factors with Overall Performance in the Simulator

Of the five CogScreen predictors, Speed/WM was found to correlate most strongly with overall flight performance, as measured by the flight summary score ($r_{rho} = 0.57$, $p < 0.0001$), accounting for 33% of the variance. Using stepwise-forward multiple linear regression analysis and the five CogScreen variables to predict flight summary scores, we found that four cognitive variables could explain 45% of the variance, F (4,77) = 16.00, $p < 0.0001$. Significant predictors, in order of entry, were Speed/WM, Visual Associative Memory, Tracking, and Motor Coordination ($p < 0.05$). This regression model is summarized in **Table IV**. The fact that Attribute Identification was not a significant predictor may be accounted for by its correlation with both Speed/WM and Tracking (age-partialled r_{rho} = 0.36 and −0.28, respectively).

We also examined a model that included age as a predictor. Previously, we found that age alone accounted for 18% of variance in flight summary scores (28). Hierarchical linear regression analysis indicated that age significantly improved prediction beyond what could be predicted by the four CogScreen variables: incremental $R^2 = 0.06$, F (1,81) = 9.55, $p = 0.003$.

Correlation of CogScreen-AE Variables with Five Indices of Simulator Performance

Table V shows the correlations of the reduced set of CogScreen-AE factors with the individual components of simulator performance. Speed/WM scores were significantly associated with 4 of the 5 indices of simulator performance: course, communication frequencies, traffic avoidance, and approach, with r_{rho} ranging from 0.39 to 0.53. Speed/WM did not predict cockpit monitoring performance ($r_{rho} = 0.06$); nor did any other CogScreen factor (r_{rho} between 0 and 0.10).

Next, we examined four relationships we had predicted between specific flight component scores and specific CogScreen scores, and obtained mixed results. First, as expected, accuracy in dialing in ATC communication frequencies was correlated with CogScreen-AE Backward Digit Span accuracy, $r_{rho} = 0.49$, $p < 0.0001$. This relationship was not simply attributable to speed of processing, as the correlation with Digit Span remained significant when the effect of speed/WM was removed, partial $r_{rho} = 0.41$, $p < 0.0001$. Second, scores on the traffic-avoidance task were but weakly to moderately correlated with spatial orientation thruput scores (Manikin task), $r_{rho} = 0.28$, $p < 0.005$, and, contrary to our expectation, this measure of spatial orienting ability did not explain variance in traffic avoidance performance beyond what was explained by general speed/WM measure, speed-adjusted $r_{rho} = -0.02$. Third, contrary to our prediction, the Visual Scanning/Sequencing factor (11), a component of our speed/WM

TABLE IV. SUMMARY OF STEPWISE FORWARD REGRESSION: COGSCREEN-AE FACTORS THAT PREDICT OVERALL FLIGHT SIMULATOR PERFORMANCE.

Step	Variable	Cum. R^2	B	± SE	F(1,81)	p
	Intercept		0.019	± 0.039	0.23	0.63
1	General Speed/WM	0.33	0.31	± 0.069	20.62	0.0001
2	Visual Associative Memory	0.39	0.144	± 0.055	6.83	0.011
3	Tracking	0.42	−0.115	± 0.049	5.46	0.022
4	Motor Coordination	0.45	0.093	± 0.043	4.73	0.033

TABLE V. CORRELATIONS OF COGSCREEN-AE FACTORS WITH FIVE INDICES OF FLIGHT PERFORMANCE AND WITH OVERALL FLIGHT SIMULATOR PERFORMANCE.

Flight Measure	General Speed/WM	Attribute Identification	Tracking	Motor Coordination	Visual Associative Memory
Course	0.49	0.23	−0.30	0.25	0.26
Communication Frequencies	0.39	0.43	−0.19	0.08	0.35
Traffic Avoidance	0.53	0.26	−0.32	0.08	0.13
Cockpit Monitoring	0.06	0.07	−0.00	0.10	0.03
Approach	0.47	0.31	−0.44	0.19	0.28
Flight Summary Score	0.57	0.41	−0.39	0.21	0.32

Notes: For flight simulator scores, higher scores indicate better performance. Except for Tracking, higher CogScreen-AE scores indicate better performance.
Spearman r_{rho} correlation coefficients. $r_{rho} \geq 0.21$ is significant at $p < 0.05$.

measure, was not significantly correlated with cockpit monitoring, $r_{rho} = 0.08$. Finally, as expected, approach scores were predicted by CogScreen-AE tracking errors, $r_{rho} = -0.44$, $p < 0.0001$. After removing variance explained by speed/WM, tracking errors remained a significant correlate of approach performance, partial $r_{rho} = -0.32$, $p < 0.002$.

DISCUSSION

These results support the validity of CogScreen-AE, illustrate the non-uniformity of age-related differences in cognitive and motor performance, and point to some areas of assessment that warrant further development. We discuss each topic in turn.

Relations of CogScreen-AE and Flight Performance

While the FAA's intent was to have a test that is predictive of brain dysfunction, our findings and those of others provide support for CogScreen-AE's occupational validity. Using multiple regression, we found that four CogScreen predictors could explain 45% of the variance in overall flight simulator performance, similar to the percentage of variance explained in a field study by Yakimovich and colleagues (26). The most powerful predictor of performance in this study was a composite score of five intercorrelated CogScreen factors that may be interpreted as general cognitive speed and working memory efficiency. This measure explained 33% of the variance in overall flight performance. Three other CogScreen factors—Visual Associative Memory, Motor Coordination, and Tracking—were found to further improve the prediction of flight simulator performance.

Our findings reinforce those of two previous studies examining relations of CogScreen scores with flight performance. In a study of 115 U.S. airline First Officers (9), five CogScreen subtest scores were correlated ($p < 0.10$) with line-check airmen ratings of how well the pilot controlled the aircraft. The tests yielding significant prediction were Visual Sequencing, Symbol Digit Coding, Manikin, Pathfinder, and Shifting Attention. The first four tests are constituents of our Speed/WM factor and the last test was designed to assess perceptual speed and executive function (11). In a second study of flight performance, a joint Russian-American field study of 75 Aeroflot captains (26), nine measures of the CogScreen battery were significant predictors of the likelihood of flight parameter violations ($p < 0.01$, $r = -0.23$ through -0.51). Of the nine measures, five were components of our speed/WM predictor. These included Shifting Attention Arrow Direction thruput (related to Visual Scanning), Matching-to-Sample speed (Visual-Perceptual/Spatial Processing), visual monitoring speed during the Divided-Attention subtest (Choice Visual Reaction time), and Previous Number task thruput in Alone and Dual conditions (Working Memory). The other significant correlates in the Russian sample were Shifting Attention Discovery thruput (Attribute Identification) and Backward Digit Span accuracy. While the latter variables did correlate significantly with flight performance in the present study, they are not elements of the stepwise multiple-regression model reported here.

The CogScreen composite measure of general cognitive speed and working memory was significantly associated with most, but not all, of the component flight scores. This, we feel, underscores the relevance of speed and working memory to aviation performance. We also found evidence for specific ability relations between certain CogScreen variables and certain simulated flight tasks. First, pilots' accuracy in dialing in ATC communication frequencies (without access to manual aids) was moderately related to speed-adjusted Backward Digit Span accuracy, which is consistent with our previous work (23). Second, we had expected that CogScreen tracking performance would predict approach scores, and this was indeed the case. Contrary to our predictions, cockpit monitoring performance was not found to be associated with any of the CogScreen predictors tested in the study and pilots' proficiency in avoiding conflicting traffic appeared to be dependent on overall speed of processing, rather than on visual-spatial skills per se. It is possible, however, that the CogScreen Manikin speed score is a conglomerate of multiple component processes—such as visual encoding, mental rotation, and response time—which, if disentangled, might provide better measures of spatial processing (5).

Age Differences in Performance of Aviation-Relevant Tasks

It appears that both age-dependent and age-independent cognitive abilities predicted performance of the simulated flight tasks in this study. Most strongly age dependent were the CogScreen-AE factors relating to cognitive speed, associative memory, and concept formation, with age-performance correlation coefficients in the range of 0.35. Motor coordination and tracking appeared to be the least age dependent, with coefficients less than 0.15. Previous findings concerning age differences in tracking performance of pilots are conflicting. Braune and Wickens reported a correlation close to zero (3), whereas two other studies found younger pilots to perform more accurately (10,25). It is worth noting that Braune and Wickens' study, designed for military aviator selection, had only 3 pilots over the age of 50 (range 20–60), whereas the latter, civilian-related studies included pilots in their 60s and 70s. The fact that the present study excluded pilots under the age of 50 may account in part for the smaller, nonsignificant correlation with age.

The age-cognition relations observed here need replication and extension, as there are few studies employing a broad set of measures over a wide age range and for specific pilot populations (8). We expect that, regardless of a sample's age span or the level of flight experience, there may be a similar pattern of age-related differences in the performance of various cognitive and motor tasks. However, the size of the age effect for a particular measure may well depend on the age span studied. Judging from the conflicting findings on age differences in tracking performance, we expect that the age-effect sizes might be smaller in the 30- to 50-yr age range than in the 50- to 70-yr age range. Indications of accelerated decline across age groups can be seen in two other studies of pilot performance (25,28). On the other hand, given evidence that flight experience is associated with better tracking performance (3,25), there may be means of mitigating the impact of age in some respects.

On the Individualized Assessment of Aviators

The data of this study and previously published studies (9,10,26) support the validity of CogScreen-AE as a cognitive battery that taps skills relevant to piloting. The existing data justify, in our opinion, further validation studies of CogScreen-AE as a clinical instrument for assessing aviators. Toward the objective of predicting aviator performance on a more individualized basis, Hyland et al. (10) and Hardy and Parasuraman (8) offer a schematic for organizing various predictors of flight performance. Briefly, the predictors can be classified as: a) domain-independent cognitive and motor skills, such as those assessable with CogScreen; b) domain-dependent (aviation) knowledge; c) pilot characteristics such as age, cardiovascular status, drug dependency, agreeableness; and d) stressors, such as difficult flight conditions, fatigue, and interpersonal conflicts. A study involving recently hired airline pilots (9) illustrates how the predictive power of these different variables can depend on the criterion or outcome of interest. For instance, domain-dependent knowledge predicted training success ($p < 0.05$) but not line-check ratings of aircraft control; conversely, CogScreen Manikin and Symbol Digit scores predicted aircraft control, but not training success. Interestingly, CogScreen Dual Task and Divided Attention scores predicted training, compliance with procedures, and crew resource management ratings, but not aircraft control. Thus, it seems unlikely that a single cognitive measure will predict aviator performance, nor is it likely that a single criterion of occupational worthiness will suffice.

In regards to the Age 60 rule, a multi-dimensional risk assessment procedure that entails both physical and cognitive examinations would most likely need to be developed in order to relax the retirement rule (e.g., (21)). Our prediction model, for example, indicated that age still accounted for a significant amount of variance beyond that predicted by CogScreen performance. This signals the possibility that other age-associated pilot characteristics might predict flight performance or carry risk for in-flight incapacitation. There is also the question of whether a given level of CogScreen performance indicates a risk for future decline in flight performance. This problem can only be answered by longitudinal study. The pilots in this project will be followed for at least 3 yr to begin addressing this question.

ACKNOWLEDGMENTS

This research was supported in part by a grant from the National Institute on Aging (AG12713), Federal Aviation Administration contract, and by the Medical Research Service of the Department of Veterans Affairs, including the Sierra-Pacific Mental Illness Research, Education, and Clinical Center (MIRECC). We would like to acknowledge Dr. Jerry Hordinsky for his contribution to research on and development of CogScreen-AE and for his enduring commitment to aviation safety. We also thank Anne Lademan and Tiffany Doelger for recruiting and testing subjects, and Allyson Rosen for critical comments on a draft of this paper.

REFERENCES

1. Banich MT, Stokes A, Elledge VC. Neuropsychological screening of aviators: a review. Aviat Space Environ Med 1989; 60:631–6.
2. Braune R, Wickens CD. Individual differences and age-related performance assessment in aviators. Part 1: Battery development and assessment. Champaign, IL: University of Illinois, Engineering Psychology Laboratory, Department of Psychology, 1984. Final Technical Report EPL-83–4/NAMRL-83–1.
3. Braune R, Wickens CD. Individual differences and age-related performance assessment in aviators. Part 2: Initial battery validation. Champaign, IL: University of Illinois, Engineering Psychology Laboratory, Department of Psychology, 1984. Final Technical Report EPL-83–7/NAMRL-83–2.
4. Callister JD, King RE, Retzlaff PD. Cognitive assessment of USAF pilot training candidates. Aviat Space Environ Med 1996; 67: 1124–9.
5. Cohen G, Faulkner D. Age differences in performance on two information-processing tasks: strategy selection and processing efficiency. J Gerontol 1983; 38:447–54.
6. Dror IE, Kosslyn SM, Wang WL. Visuospatial abilities in pilots. J Applied Psych 1993; 5:763–73.
7. Gordon HW, Leighty R. Importance of specialized cognitive function in the selection of military pilots. J Appl Psychol 1988; 73:38–45.
8. Hardy DJ, Parasuraman R. Cognition and flight performance in older pilots. J Exp Psychol (Appl) 1997; 3:313–48.
9. Hoffmann CC, Hoffmann KP, Kay GG. The role that cognitive ability plays in CRM. Paper presented at NATO Symposium, Human Factors & Medicine Panel on Collaborative Crew Performance in Complex Operational Systems, Edinburgh, Scotland. April 1998.
10. Hyland DT, Kay EJ, Deimler JD. Age 60 study, Part IV: Experimental evaluation of pilot performance. Washington, DC: Office of Aviation Medicine 1994; DOT/FAA/AM-94/23.
11. Kay GG. CogScreen aeromedical edition professional manual. Odessa, FL: Psychological Assessment Resources, Inc., 1995.
12. King RE, Flynn CF. Defining and measuring the "right stuff": neuropsychiatrically enhanced flight screening (N-EFS). Aviat Space Environ Med 1995; 66:951–6.

13. Leirer VO, Yesavage JA, Morrow DG. Marijuana carry-over effects on aircraft pilot performance. Aviat Space Environ Med 1991; 62:221–7.
14. Lindenberger U, Mayr U, Kliegl R. Speed and intelligence in old age. Psychol Aging 1993; 8:207–20.
15. Morrow D, et al. The time-course of alcohol impairment of general aviation pilot performance in a Frasca 141 simulator. Aviat Space Environ Med 1993; 64:697–705.
16. Morrow D, et al. The influence of aging and practice on piloting tasks. Exp Aging Res 1993; 19:53–70.
17. Mumenthaler MS, et al. Influence of nicotine on simulator flight performance in non-smokers. Psychopharmacology 1998; 140: 38–41.
18. North RA, Gopher D. Measures of attention as predictors of flight performance. Hum Fact 1976; 18:1–14.
19. Salthouse TA. Mediation of adult age differences in cognition by reductions in working memory and speed of processing. Psychol Sci 1991; 2:179–83.
20. Salthouse TA. The processing-speed theory of adult age differences in cognition. Psychol Rev 1996; 103:403–28.
21. Stuck AE, et al. Multidimensional risk assessment versus age as criterion for retirement of airline pilots. J Am Geriatr Soc 1992; 40:526–32.
22. Taylor JL, et al. Acute and 8-hour effects of alcohol (0.08% BAC) on younger and older aircraft pilots' simulator performance. Aviat Space Environ Med 1994; 65:718–25.
23. Taylor JL, et al. The effects of information load and speach rate on younger and older aircraft pilots' ability to execute air-traffic controller instructions. J Gerontol B-Psychol 1994; 49:P191–P200.
24. Tsang PS. Can pilots time-share better than non-pilots? Appl Ergon 1986; 17:284–90.
25. Tsang PS. Boundaries of cognitive performance as a function of age and flight experience. Int J Aviat Psychol 1996; 6:359–77.
26. Yakimovich NV, et al. CogScreen as a predictor of flight performance in Russian pilots. Paper presented at 65th annual scientific meeting of the Aerospace Medical Association, San Antonio, TX. May, 1994.
27. Yesavage JA, Dolhert N, Taylor JL. Flight simulator performance of younger and older aircraft pilots: Effects of age and alcohol. J Am Geriatr Soc 1994; 42:577–82.
28. Yesavage JA, et al. Relationship of age and simulated flight performance. J Am Geriatr Soc 1999; 47:819–23.