Cognitive Evaluation of Israeli Air Force Pilot Cadets

Shirley Gordon; Chen Goren; Erez Carmon; Leah Shelef

OBJECTIVE: In aviation psychology, there is a constant need for the cognitive evaluation of pilots as part of operational fitness and safety criteria.

METHODS: A cross-sectional study with comparison between the performance of Israeli Air Force pilot cadets ($N = 318$) and U.S. Air Force pilot training candidates ($N = 512$) as assessed by a cognitive battery was undertaken. The data of the comparison group was collected from Callister, King, and Retzlaff, as published in 1996.

RESULTS: In general, the means in the three components composing the battery—speed, accuracy, and throughput variables—indicated that the Israeli Air Force pilot cadets' scores were higher than those of the U.S. Air Force pilot candidates' scores in 50 of 53 variables. Nonsignificant differences were found in Accuracy of shifting attention-arrow color (SATAC), pathfinder-combined (PFC), and pathfinder-letter (PFL).

CONCLUSIONS: The difference in performance between the two groups may be due to differences in population characteristics. However, these results need to be considered cautiously, as the groups were sampled at a sizeable time gap (1996 for the U.S. Air Force vs. 2013 for the Israeli Air Force), with each time period characterized by different cultural and technological influences.

KEYWORDS: aviation psychology, pilot cadets, cognitive evaluation.

Gordon S, Goren C, Carmon E, Shelef L. *Cognitive evaluation of Israeli Air Force pilot cadets*. Aerosp Med Hum Perform. 2017; 88(4):392–398.

Flying a jet is a complex task, requiring very specific cognitive abilities and personality traits. Selecting the most adequate individuals to serve as pilots involves a particular challenge and, for that reason, the selection and evaluation process is prolonged and demanding. Moreover, since the training procedure and the acquisition of this specific profession is so valuable and resource consuming, it is important to accurately diagnose a decline in performance or in cognitive abilities resulting from medical conditions.[3]

In order to map an individual's cognitive functions, a comprehensive psychological evaluation is required, comprising assessment of intelligence, attentional functions, memory, information processing, spatial, and executive function abilities.[16] Recognizing the need for the cognitive evaluation of pilots,[3] the U.S. Federal Aviation Administration (FAA) sponsored the development of a cognitive battery specific for aviators.[8]

The battery was designed to measure numerous abilities associated with flying and related to flying performance.[4,7] It consists of a series of 11 computerized cognitive tasks, each self-contained and presented with instructions and a practice segment, yielding a total of 65 scores. The 11 tasks comprise background digit span, math, visual sequence comparison, symbol digit coding, matching to sample, manikin, divided attention, auditory sequence comparison, pathfinder, shifting attention, and dual task. Each of the tasks can be scored in several ways. Typical scoring includes task speed (RTC suffix), accuracy (ACC suffix), and throughput (PUT suffix), which is the number of correct reaction responses per minute.[4] These three variables, arising from a given assessment, are very likely to be collinear, due to the arithmetic relationship between speed (reaction time) and throughput (especially applicable when accuracy rates approach 100%).[4]

According to Kay,[8] the battery can rapidly assess deficits or changes in attention, immediate and short-term memory, visual-perceptual functions, sequencing functions, logical problem solving, calculation skills, reaction time, simultaneous

From the Israeli Air Force and IDF Medical Corps, Tel Hashomer, Israel.

This manuscript was received for review in May 2016. It was accepted for publication in January 2017.

Address correspondence to: Shirley Gordon, M.A., Psychology Branch, Chief Surgeon Headquarters, Israeli Air Force, Israel Defense Forces (IDF), Tel Hashomer, Ramat-Gan, Israel; Shirley.gordon@gmail.com.

DOI: https://doi.org/10.3357/AMHP.4677.2017

information processing abilities, and executive functions. Moreover, the cognitive battery variables were able to explain 45% of the variance of flight performance, as measured by a flight simulator,[14,18] and hence found to be suitable for pilot assessment.[4] It was also found to be a sensitive tool for discriminating between neurologically impaired and cognitively intact individuals.[8]

Callister et al.[4] compared USAF pilot candidate evaluations to commercial pilot norms and found differences between the two populations. While the former was more accurate on measures of numeric working memory, completed divided attention and other dual tasks faster, and were more impulsive, the latter were more accurate in solving arithmetic word problems, completed simple and focused tasks faster, and were more perseverative. The authors explained the differences between the two populations by their age difference, a variable found to be highly correlated with cognitive abilities, as the USAF pilot candidates were much younger ($M_{age} = 23.5$) than the commercial pilots ($M_{age} = 44$). Indeed, Taylor et al.[14] examined the relationship of the cognitive battery to flight simulator performance and pilot age, finding that incorporating the age variable significantly enhances performance prediction. In addition, they demonstrated that speed\working memory, visual associative memory, motor coordination, and tracking were significant predictors of pilot performance. Even though this cognitive battery is one of the few neuropsychological test batteries with comprehensive pilot norms, research on it is still scant.[9]

Since this battery was identified as an appropriate device for the assessment of pilots, it has been used worldwide in aviation medicine decision making, with the air forces of several countries using it as a neurocognitive assessment. It was used by the U.S. Air Force (USAF) as a part of the USAF Enhanced Flight Screening program;[13] however, in recent years, the USAF has integrated a new computerized assessment battery called MicroCog.[14]

National military service in Israel is mandatory for most Israeli citizens reaching the age of 18. In advance of their enlistment, Israeli recruits undergo a series of tests and evaluations to determine their suitability for military service, thus resulting in their being mentally healthier as a group than the general population.[5] Approximately 60% of men and 40% of women are typically found fit to serve. Among those found suitable for general army duty, only those with the highest intellectual and psychomotor abilities can be referred for an evaluation to join the pilot cadets' academy. For a subpopulation with high cognitive abilities and low variance in testing scores, it is much more difficult to detect a decrease in performance when compared to the general population.[9]

Understanding the problematic use of conventional neuropsychological tests for assessment of pilots, the Israeli Air Force (IAF) began only recently to use this battery (Hebrew Edition) for all pilot cadets reaching the late stages of the pilot cadets' academy. This test is one of the final routine medical examinations a pilot is required to take prior to approving his/her medical certification for flying and, by the time of taking this test, Israeli pilot cadets had approximately 100 h of flying instruction on their resume. It is important to note that this assessment is not used for selection, but is rather used as a diagnostic tool for future neuropsychological evaluation. It is aimed at creating a personal profile of cognitive function for each young pilot so that, in a case of neurological impairment, an accurate index of premorbid cognitive performance will be available for comparison and for evaluation of his\her fitness to return to duty. Currently, only the pilot cadet graduates take the test in light of cost-benefit considerations.

This cognitive battery is widely used as a neuropsychological tool to assess pilot's cognitive performance. Yet current pilot psychometric norms are limited and are usually based on small, specialized populations. As comparisons between different pilot populations are rare,[9] this paper seeks to contribute to the existing literature by comparing the cognitive scores of IAF pilot cadets to those of USAF pilot training candidates. Moreover, on the basis of a previous comparison between young USAF pilots and older commercial pilots, having yielded differences in cognitive performance based on age differences,[4] it would be valuable to examine cognitive performance in two populations with similar age and other characteristics. To the best of our knowledge, this is the first study conducted in the IAF that compares between populations (IAF and USAF) based upon the cognitive testing protocol.

METHODS

A cross-sectional study[15] using the CogScreen-AE (CogScreen, LLC, St. Petersburg, FL; http://www.cogscreen.com) scores of both IAF pilot cadets and USAF pilot training candidates from different time periods (2011–2014 and 1996, respectively) was undertaken. All data were collected via computerized assessment software (Windows XP based operating system with Intel Core-Duo microprocessor[1]), which administered the test, timed the tasks, scored the test, and archived the data in report form. As noted, this is one of the final routine medical examinations required of pilots prior to certification.

Subjects

A sample of 318 IAF pilot cadets participated in this study, with a mean age of 22.05 yr ($SD_{age} = 1.32$; ranging from 20.3 to 29.8). Of the cadets, 94% were men and all were at the final stage of the pilot cadets' academy, and after college graduation. The institutional review board of the IDF Medical Corps approved the study and waived the requirement for informed consent in the interest of preserving subjects' anonymity. The data of a comparison group collected from Callister, King, and Retzlaff in 1996[4] consisted of 512 USAF pilot training candidates, with a mean age of 23.5 ($SD_{age} = 4.2$). Of these, 95% were men and approximately 42 were juniors at the USAF Academy, the remainder having been commissioned through Officer Training School. Reserve Officer Trainees and the Air National Guard subjects were all college graduates. This hardware and software combination was confirmed as a valid platform for administration and would allow scores to be compared between 1996 and the current study.

Materials

The battery was used as instructed by the test publisher. Further details can be found in the manual.[8] The total test takes about 50 min to complete. The program runs on a standard computer, attached to a touchscreen operated with a plastic pen. The plastic pen allows the respondent to maintain focus on the screen at all times. The battery takes advantage of the many benefits offered by computerization, including improved standardization of administration and scoring, superior time recording, and the presentation of dynamic and multiple test stimuli. The program automatically generates 12 alternate test forms, which may be used for repeated test administrations for the same individual. The program presents each test item and automatically records the response and the response time.

Within the 11 tasks of the battery,[8] only measures common to both the USAF pilot training candidates and the IAF pilot cadets were included in this study. Thus, 11 tasks were able to be compared and are presented in **Table I**.

Table I. CogScreen-AE Tasks.

TASKS	DESCRIPTION
Backward digit span (BDS)	Numbers presented one at a time. The task is to remember these numbers and, when indicated, enter the numbers in reverse order using the pointer. There is no time limit on this task.
Math (MATH)	Traditional math word problems with multiple-choice answer format.
Visual sequence comparison (VSC)	Comparison of two simultaneously presented series of letters and numbers. The respondent determines whether two alphanumeric strings presented side by side are the same or different.
Matching to sample (MTS)	Immediate recognition for a checkerboard pattern: after viewing a four-by-four grid pattern, the respondent is asked to select the correct pattern from two grids displayed side by side.
Manikin (MAN)	Mental-rotation task, requiring respondent to identify the hand in which a rotated human figure is holding a flag (one of four positions).
Divided attention (DAT, DATI, DATD, DATSC)	Monitor the vertical movement of a cursor within a circle and return it to center when it exceeds the boundaries. The task is performed alone and with the visual sequence comparison task.
Auditory sequence comparison (ASC)	Comparison of two series of four to eight tones of varying pitch, presented sequentially.
Symbol digit coding (SDCACC, SDCIRACC, SDCDRACC)	Measures of attention, visual scanning, working memory, and speed of information processing.
Pathfinder (PF, PFN, PFL, PFC)	Visual sequencing and scanning task that requires respondents to sequence numbers, letters, and an alternating set of numbers and letters.
Shifting attention (SAT, SATAD, SATAC, SATIN, SATDI)	Rule-acquisition and rule-application test requiring mental flexibility and conceptual reasoning.
Dual task (DTT, DTTPA, DTTPD)	Consists of two tasks, each of which is performed alone and then together as a simultaneous test: 1) a visual-motor tracking test; and 2) a continuous memory task involving serial recall.

Backward digit span. In backward digit span (BDS), random numbers are presented one at a time. The task is to remember these numbers and, when indicated, to enter the numbers in reverse order using the pointer. There is no time limit on this task.

Math. The math (MATH) task presents five traditional multistep math problems (at approximately the 10th-grade level of difficulty). After reading the test item, the subject is asked to select a response from among three options. Performance on the MATH subtest requires computational math skills, attention, concentration, working memory, reading comprehension, and logical reasoning.

Visual sequence comparison. The visual sequence comparison (VSC) task simultaneously presents the respondent with two alphanumeric strings, one on the right half of the screen and the other on the left half. The respondent selects SAME or DIFFERENT for each pair, indicating whether the same characters are pre-entered in the same order for both sequences. The strings vary in length from four to eight items. For each pair, the strings may vary by one or two numbers or letters. Half of the 20 sequence pairs present a congruent sequence, and half present an incongruent sequence. This subtest includes visual attention, working memory, verbal sequential processing, and visual-perceptual speed.

Matching to sample. In the matching to sample (MTS) task, one pattern is the same as the original pattern, and the other pattern differs in 1 of the 16 cells. The respondent's task is to select the grid corresponding to the original grid pattern. There are 20 trials of this task presented, with approximately half of the matching patterns presented on the right and half on the left sides of the screen. MTS measures visual-perceptual speed, spatial processing, and visual working memory.

Manikin. The manikin (MAN) task presents a male human figure holding a flag which is either right-side up, upside-down, facing forward, or facing backward. The respondent's task is to determine whether the figure is holding the flag in his left or right hand. Twenty trials are presented with an equal left- and right-hand presentation. MAN measures visual-spatial perception, spatial orientation, and the ability to mentally rotate visual images.

Divided attention. The divided attention (DAT) task consists of three components: 1) in the visual monitoring only task (DATI), the respondent watches a cursor move vertically within a circle that is divided into central, upper, and lower sections. When the cursor crosses into the upper or lower section, the respondent is required to press a box marked "center" with the light pen. The amount of time the indicator remains in the upper and lower section of the circle prior to being centered by the respondent is also measured. 2) In simultaneous (dual) presentation of the visual monitoring task (DATD), the response time is measured for both the visual monitoring task and the VSC. When the two tasks are presented simultaneously, the test assesses divided attention, working memory, and visual-motor and visual-perceptual speed. A comparison of performance under single and dual task conditions for sequence comparison, visual monitoring of the indicator, and premature responses yields information regarding the respondent's capacity for multitasking. 3) The visual sequence comparison task (DATSC), as noted above (VSC), comprises the third component.

Auditory sequence comparison. The auditory sequence comparison (ASC) task requires comparison of two series of four to eight tones. The series, presented sequentially, may differ in the pitch of one tone. The respondent indicates SAME or DIFFERENT for each sequence pair. There are 10 auditory-sequence pairs presented, with half of the sequences the same and the other half different. Performance on ASC requires auditory attention, working memory, and sound pattern discrimination.

Symbol digit coding. Symbol digit coding (SDC) is a computer analogue of the conventional symbol-substitution task found in the WAIS-R Digit symbol subtest.[12] Six symbol-digit pairs are displayed continuously throughout the test near the top of the screen. The respondent is instructed to remember the symbol-digit pairs for a subsequent memory test. In the center of the screen, a line of symbols in random order is presented with blank spaces directly beneath each symbol. The respondent selects the associated digit for each symbol. As each row is completed, a new row appears. Performance measures include response accuracy (SDC Accuracy: SDCACC) and the number of correct responses per minute (SDC Throughput; SDCPUT). This subtest measures attention, visual scanning, working memory, and speed of information processing. The SDC immediate recall task (SDCIRACC) is an immediate recall trial of the symbol-digit pairs following the 90-s trial of the SDC task. The SDC delayed recall Task (SDCDRACC) is presented after the pathfinder subtest, following a 20- to 30-min delay and is identical to SDCIRACC.

Pathfinder. In the pathfinder (PF) task, the respondent is required to select one of the four quadrants containing the next character of a previously specified sequence. There are three different sequencing rules: PF number requires number sequencing (PFN), PF letter requires letter sequencing (PFL), and PF combined (PFC) requires sequencing an alternating series of numbers and letters. PF measures number and letter sequencing skills, the ability to systematically apply an organizing principle, immediate memory, motor coordination, visual scanning, and the ability to shift mental set.

Shifting attention. In the shifting attention (SAT) task, respondents are required to alter their responses depending upon changing rules. SAT begins by training the respondent to select from among four response boxes, each of which contains an arrow, according to one of three easily learned rules: select a box based on the color of its border; select a box based on the direction of the arrow (SATAD); and match the color of the arrow (SATAC). After learning the three response conditions, the respondent begins the fourth condition, in which he is requested to identify the active rule before the presentation of each subsequent stimulus (SATIN). In the fifth condition, the respondent's task is to discover and then apply the active response rule, which changes after a variable number of correct responses (SATDI). SAT is a test of concept formation (attribute identification), mental flexibility, sustained attention, deductive reasoning (rule-learning), vulnerability to response interference, working memory, application of novel rules, visual scanning, choice visual reaction time, and perseverative tendencies.

Dual task. In the dual task (DTT), two tasks are each presented independently and then simultaneously for a measure of divided attention and multitasking capability. In the independent task (DTTPA), the respondent is required to recall the previous number shown and to select that number, while simultaneously encoding the current number for the next stimulus presentation; in the simultaneous task (DTTPD), the tracking and the previous number tasks are presented simultaneously. DTT measures sustained attention, visual-motor tracking, divided attention, and working memory.

In addition to the 11 cognitive tasks described above, three scoring methods were also carried out:

Speed. Speed measures are the median reaction time on correct trials of battery tasks. The median, rather than the mean, is used because the median reduces the effect of outlier response times.

Accuracy. Accuracy is the percentage of correct responses to the total number of administered items: number of correct responses/total number of items $\times$ 100. Total number of items includes correct responses, incorrect responses, and lapses (i.e., failures to respond within time limits).

Throughput. Throughput is a function of speed and accuracy, basically the number of correct reaction responses per unit of time (minute). Throughput = [(Accuracy/100) $\times$ 60 s]/median response time for correct trials.

Statistical Analysis

Analyses were carried out with IBM SPSS, version 20. As mentioned above, the USAF pilot candidate group data were presented by Callister et al.[4] only as mean scores with standard deviations (no individual raw data). Therefore, only t-tests could be conducted. While multivariate statistics would have been preferable, these would have required the original data set. As the current study was exploratory in nature, with no hypotheses regarding the two similar groups, two-tailed, independent *t*-tests were calculated for each of the study variables.

RESULTS

IAF and USAF groups were first compared in terms of age and gender, as those were the only comparable demographic variables, based on the USAF pilot candidates' given data.[4] In order to maintain congruence with the Callister et al.[4] study, the results of the battery variables are presented by subtest and by type of score so that speed variables are presented first, followed by accuracy, and then throughput.

By comparing the means of the speed (**Table II**), accuracy (**Table III**), and throughput (**Table IV**) variables of IAF pilot cadets and USAF pilot candidates, it was found that the IAF pilot cadets achieved higher scores in all measures [$P < 0.001$; except for the variables of: Accuracy in shifting attention-arrow color (SATAC), pathfinder-combined (PFC), and pathfinder-letter (PFL)].

Table III. Means and Standard Deviations With *t*-Tests for IAF Pilot Cadets (N = 318) and USAF Pilot Candidates (N = 512) on CogScreen-AE Accuracy Variables.

VARIABLE	IAF M (SD)		USAF* M (SD)		DF	*t*-TEST	*P*-VALUE (2-TAIL)
BDSACC	0.87	(0.14)	0.89	(0.12)	828	1.70	0.022
MATHACC	0.84	(0.19)	0.72	(0.19)	828	−8.94	<0.001
VSCACC	0.98	(0.03)	0.97	(0.03)	828	−6.13	<0.001
SDCACC	0.99	(0.01)	0.99	(0.01)	828	−5.47	<0.001
SDCIRACC	0.99	(0.06)	0.94	(0.13)	828	−6.40	<0.001
MTSACC	0.97	(0.04)	0.95	(0.05)	828	−7.39	<0.001
MANACC	0.95	(0.09)	0.93	(0.09)	828	−3.59	<0.001
DATSCACC	0.93	(0.05)	0.89	(0.07)	828	−8.82	<0.001
ASCACC	0.95	(0.07)	0.90	(0.10)	828	−8.44	<0.001
PFNACC	1.00	(0.01)	0.99	(0.01)	828	−11.60	<0.001
PFLACC	0.98	(0.04)	0.99	(0.01)	828	0.55	0.145
PFCACC	0.98	(0.04)	0.98	(0.03)	828	0.89	0.094
SDCDRACC	0.99	(0.07)	0.93	(0.15)	828	−6.12	<0.001
SATADACC	0.98	(0.05)	0.98	(0.03)	828	−1.77	0.019
SATACACC	0.99	(0.03)	0.99	(0.03)	828	−0.02	0.246
SATINACC	0.98	(0.03)	0.97	(0.03)	828	−3.67	<0.001
SATDIACC	0.74	(0.07)	0.67	(0.11)	828	−10.85	<0.001
DTTPAACC	0.97	(0.03)	0.93	(0.07)	827	−9.36	<0.001
DTTPDACC	0.95	(0.04)	0.86	(0.11)	827	−14.02	<0.001

* USAF data abstracted from Callister et al. (1996).[4]

DISCUSSION

This study compared IAF pilot cadets to USAF pilot candidates in light of the two populations' similarity in age and other characteristics. Nevertheless, differences in most cognitive measurements were found between the two groups. The difference in the mean age of the two groups (22.05 and 23.5, respectively) was not significant. Indeed, studies that examined age-related influences on cognitive abilities across the life span required a broader age range with larger intervals in order to achieve statistical significance.[1,2,6] Moreover, Morrow et al.[10] examined three age ranges (22–40, 50–59, 60–76) of pilots and nonpilots in terms of expertise and cognitive and visual-spatial abilities. These researchers found that age did not moderate the association between expertise and pilot performance, thus concluding that functional age (experience) is a better predictor than chronological age of pilot performance.[11] The nonsignificant difference in age between the two groups would not seem to explain the significant differences in 50 of the 53 variables.

A second possible explanation for the noted differences in the cognitive battery scores could be cognitive capacity. Possibly the selection process for the IAF pilot cadets brings a more homogenous group with higher innate cognitive abilities than

Table II. Means and Standard Deviations With *t*-Tests for IAF Pilot Cadets (N = 318) and USAF Pilot Candidates (N = 512) On CogScreen-AE Speed Variables.

VARIABLE	IAF M (SD)	USAF* M (SD)	DF	*t*-TEST	*P*-VALUE (2-TAIL)
MATHRTC	24.93 (7.77)	27.25 (8.79)	828	3.85	<0.001
VSCRTC	2.06 (0.45)	2.24 (0.51)	828	5.29	<0.001
MTSRTC	1.02 (0.20)	1.47 (0.28)	828	24.81	<0.001
MANRTC	1.54 (0.34)	1.98 (0.38)	828	16.84	<0.001
DATIRTC	0.24 (0.05)	0.40 (0.07)	828	35.97	<0.001
DATDRTC	0.35 (0.11)	0.69 (0.20)	828	27.75	<0.001
DATSCRTC	1.91 (0.63)	2.15 (0.53)	828	5.91	<0.001
ASCRTC	0.66 (0.15)	0.98 (0.24)	828	21.12	<0.001
PFNRTC	0.76 (0.22)	0.85 (0.16)	828	6.59	<0.001
PFLRTC	0.77 (0.20)	0.79 (0.13)	828	1.44	0.037
PFCRTC	0.92 (0.22)	1.20 (0.30)	828	14.33	<0.001
SATADRTC	0.49 (0.08)	0.70 (0.10)	828	31.59	<0.001
SATACRTC	0.52 (0.07)	0.68 (0.09)	828	26.00	<0.001
SATINRTC	0.61 (0.11)	0.86 (0.15)	828	26.15	<0.001
SATDIRTC	0.73 (0.15)	0.95 (0.21)	828	16.45	<0.001
DTTPARTC	0.34 (0.14)	0.48 (0.19)	826	11.52	<0.001
DTTDRTC	0.43 (0.14)	0.66 (0.24)	826	150.61	<0.001

* USAF data abstracted from Callister et al. (1996).[4]

Table IV. Means and Standard Deviations With *t*-Tests for IAF Pilot Cadets (N = 318) and USAF Pilot Candidates (N = 512) on CogScreen-AE Throughput Variables.

VARIABLE	IAF M (SD)		USAF* M (SD)		DF	*t*-TEST	*P*- VALUE (2-TAIL)
MATHPUT	2.27	(0.95)	1.82	(1.22)	828	−5.61	<0.001
VSCPUT	30.14	(6.77)	27.56	(6.20)	828	−5.62	<0.001
SDCPUT	55.97	(11.41)	33.74	(6.00)	828	−36.68	<0.001
MTSPUT	59.07	(10.74)	40.44	(7.73)	828	−28.98	<0.001
MANPUT	39.08	(9.54)	29.51	(7.05)	828	−16.55	<0.001
DATSCPUT	31.95	(9.04)	26.32	(6.47)	828	−10.43	<0.001
ASCPUT	90.00	(19.29)	58.79	(17.48)	828	−24.02	<0.001
PFNPUT	83.40	(19.01)	72.00	(12.86)	828	−10.30	<0.001
PFLPUT	80.15	(15.48)	77.46	(12.24)	828	−2.78	0.001
PFCPUT	66.63	(13.70)	51.83	(12.54)	828	−15.95	<0.001
SATADPUT	122.21	(18.52)	86.55	(12.77)	828	−32.79	<0.001
SATACPUT	115.14	(14.82)	88.51	(11.20)	828	−29.35	<0.001
SATINPUT	99.55	(16.57)	69.59	(11.64)	828	−30.54	<0.001
SATDIPUT	63.15	(12.67)	44.64	(11.68)	828	−21.48	<0.001
DTTPAPUT	223.21	(266.67)	131.25	(46.15)	825	−7.62	<0.001
DTTPDPUT	166.70	(190.68)	90.85	(38.48)	825	−8.72	<0.001

* USAF data abstracted from Callister et al. (1996).[4]

the self-selection process of the USAF pilot candidates. As we have no specific data on cognitive capacity by other standardized IQ tests (such as the Wechsler Adult Intelligence Scale[17]) in the two groups, this question must be left unanswered.

A third possible explanation of the data differences could be related to academic training and possible flying experience. As mentioned before, the IAF pilot cadets had approximately 100 flying training hours prior to taking the test. While both the IAF pilot cadets and the USAF pilot candidates are all "nearly-completed" or "completed" university graduates, the undergraduate courses of 1996 vs. 2013 may be significantly different. It is possible that current students are better prepared to solve the subtests of the test due to their curriculum. In addition, the practical experience and aviation knowledge acquired during the IAF pilot cadets' academy may have some positive impact on test scores. Because the USAF is now using a different battery rather than this battery for baseline cognitive testing, it is not possible to request more current normative data for USAF pilot candidates.

A fourth possible explanation is the impact of culture and technology, which has changed dramatically between 1996 and 2013. Thus, the IAF cadets' superior performance on the computerized battery may be explained by their greater familiarity with modern technology, particularly their computer literacy. Indeed, in their review, Oblinger and Oblinger[11] argued that today's youth are likely to be developing greater digital literacy than even their slightly older siblings. For example, they reported in 2005 that over two million American children (ages 6 to 17) had their own website, a phenomenon much less likely a decade earlier. The ability to use nontext expression, such as audio, video, and graphics, and their interaction, appears stronger in each successive age cohort.

From this perspective, it is possible that 2013 USAF pilot candidates would also show greater proficiency in battery subtests, as generational cohorts to the IAF pilot cadets. Given that there is no data available to define if there has been score "inflation" over the 20 yr of administering it, we cannot speak further to this possibility.

Future research is warranted to establish a better understanding of the role of preflight and academic training on the performance of pilot candidates who take the test. It would be an advantage if air forces of other countries who are currently using the test in their pilot assessments also presented mean scores and standard deviations. If large differences are noted, then the technology familiarity may not be a strong explanation for variances, but perhaps training curricula. Since higher test scores have tended to suggest higher flying skills, finding improved training curricula would be a benefit for all new pilots. Our hope is that further data on the test scores from earlier decades to present will become available in the medical literature as well. If technology familiarity is creating an increase in subtest scores, then a reconsideration of the normative pilot data may be needed.

There are a number of limitations that should be noted. One limitation relates to the study's cross-sectional design, which precludes inference of causality. Secondly, the constraint of using *t*-tests in this study may contribute to accumulated statistical error. The third limitation relates to the preliminary group differences: 1) the IAF 'cognitive capacity' could be higher than the USAF pilot candidates; 2) IAF pilot cadets had approximately 100 flying training hours prior to taking the battery; and 3) we have no data to assess whether the current pilots' computer literacy skills may have artificially elevated the scores now compared to those pilots taking the exam 20 yr ago.

This paper seeks to contribute to the relatively rare though existing literature of neuropsychological evaluation of pilots by comparing the cognitive scores of IAF pilot cadets to those of USAF pilot training candidates. The difference in performance between the two groups may be due to differences in population characteristics. However, these results need to be considered cautiously, as the groups were sampled at a sizeable time gap (1996 for the USAF vs. 2013 for the IAF), with each time period characterized by different cultural and technological influences.

ACKNOWLEDGMENTS

Shirley Gordon had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis. She is also responsible for the study concept and design. All of the authors participated in the analysis, interpretation, and statistical analysis of the data.

All authors report no financial relationships with commercial interests.

Authors and affiliation: Shirley Gordon, M.A., Chen Goren, M.A., Erez Carmon, M.D., and Leah Shelef, Ph.D., Chief Surgeon Headquarters, IDF Psychology Branch, IDF Medical Corps, Israel Air Force, Tel Hashomer, Ramat-Gan, Israel.

REFERENCES

1. Angel L, Bastin C, Genon S, Salmon E, Fay S, et al. Neural correlates of successful memory retrieval in aging: do executive functioning and task difficulty matter? Brain Res. 2016; 1631:53–71.
2. Baltes PB, Lindenberger U. Emergence of a powerful connection between sensory and cognitive functions across the adult life span: a new window to the study of cognitive aging? Psychol Aging. 1997; 12(1):12–21.
3. Banich MT, Stokes A, Elledge VC. Neuropsychological screening of aviators: a review. Aviat Space Environ Med. 1989; 60(4):361–366.
4. Callister JD, King RE, Retzlaff PD. Cognitive assessment of USAF pilot training candidates. Aviat Space Environ Med. 1996; 67(12):1124–1129.
5. Gal R. A portrait of the Israeli soldier (No. 52). NewYork: Greenwood Publishing Group; 1986.
6. Hardy DJ, Satz P, D'Elia LF, Uchiyama CL. Age-related group and individual differences in aircraft pilot cognition. Int J Aviat Psychol. 2007; 17(1):77–90.
7. Imhoff DL, Levine JM. Perceptual-motor and cognitive performance task battery for pilot selection. Brooks AFB (TX): Air Force Human Resources Laboratory; 1986.
8. Kay GG. CogScreen Aeromedical Edition professional manual. Odessa (FL): Psychological Assessment Resources; 1995.
9. King RE, Flynn CF. Defining and measuring the 'right stuff': Neuropsychiatrically-Enhanced Flight Screening (N-EFS). Aviat Space Environ Med. 1995; 66(10):951–956.
10. Morrow DG, Miller LMS, 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(1):P11–P18.
11. Oblinger D, Oblinger J. Is it age or IT: first steps toward understanding the net generation. In Oblinger D, Oblinger J, editors. Educating the net

generation. 2005; Educause e-book 2.1–2.20 [Accessed May 25, 2016]. Available from https://net.educause.edu/ir/library/pdf/pub7101d.pdf.
12. Powell DH, Kaplan EF, Whitla D, Weintraub S, Caitlin R, Funkenstein HH. MicroCog: assessment of cognitive functioning Windows edition (MicroCog for Windows). Manual/Installation Guide. San Antonio (TX): Pearson; 2004.
13. Retzlaff PD, Callister JD, King RE. Clinical procedures for the neuropsychological evaluation of U.S. Air force pilots. Mil Med. 1999; 164(7):514–519.
14. 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(4):373–380.
15. Vandenbroucke JP, von Elm E, Altman DG, Gøtzsche PC, Mulrow CD, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. Ann Intern Med. 2007; 147(8):W163–W194.
16. Vanderploeg RD, editor. Clinician's guide to neuropsychological assessment, 2nd ed. Mahwah (NJ): Lawrence Erlbaum Associates; 2000.
17. Wechsler D. Wechsler Adult Intelligence Scale-Revised. New York: The Psychological Corporation; 1981.
18. Yakimovich NV, Strongin GL, Govorushenko VV, Schroeder DJ, Kay GG. CogScreen as a predictor of flight performance in Russian pilots. Paper presented at the 65th annual scientific meeting of the Aerospace Medical Association, San Antonio, TX; 1994 May. Alexandria (VA): AsMA; 1994.