THE FINE MOTOR SKILLS AND COGNITION TEST BATTERIES: NORMATIVE DATA AND INTERDEPENDENCIES

Bettina L. Beard[(1)], Albert J. Ahumada, Jr.[(2)], Kritina L. Holden[(3)]

[(1)]NASA Ames Research Center, Moffett Field, CA 94040, U.S.A., tina.beard@nasa.gov,
[(2)]San Jose State University Foundation @ NASA Ames, U.S.A., ajahujr@gmail.com
[(3)]Wyle Laboratories @ NASA Johnson Space Center, U.S.A., kritina.l.holden@nasa.gov

ABSTRACT

Fine motor skills and cognitive abilities are major contributors to crew performance on essentially all extravehicular and intra-vehicular activities during spaceflight. It is critical for the crew's safety, and for mission productivity, to know if, and when, motor skills or cognitive abilities are compromised so that countermeasures may be introduced. NASA has developed two test batteries to measure and monitor astronaut cognitive and fine motor skills. The Cognition Test Battery contains 10 sub-tests that assess cognitive behaviors ranging from low level visual perception to high level decision-making. The Fine Motor Skills Test Battery contains 4 sub-tests that assess finger dexterity, manual dexterity and wrist-finger speed. This study sought to determine acceptable norms for both batteries in an astronaut-like population and to identify the extent to which fine motor skills contribute to cognitive test scores.

1. INTRODUCTION

Extravehicular and intra-vehicular schedules and activities are physically and mentally demanding. Astronauts and cosmonauts must perform various tasks that require visual perception, eye-hand coordination, sustained and divided attention, working and long-term memory, logic, task prioritization and decision-making in the presence of numerous and diverse spaceflight stressors. For this reason, NASA is developing two computerized test batteries to assess crew performance before, during and after missions. The Cognition Test Battery (CTB) assesses ten cognitive domains[1] while the Fine Motor Skills Test Battery (FMS) assesses three measures of manual brain function.

Cognitive and motor processes are often discussed as if they are independent functions (e.g., sensation as separate than executive processes or decision making as separate than motor responses), but they are actually inter-dependent. There are several lines of evidence suggesting an inter-dependency between fine motor skills and cognition. For example, terrestrial research shows that both cognitive and fine motor processes are disrupted by similar stressors such as workload[2,3], vibration[4,5], G-forces[6,7], danger[8], fatigue[9,10], CO_2 [11], sleep deprivation[12] and extreme environments[13].

If a crew member is showing poorer performance on any CTB or FMS sub-test, knowledge about what type of cognitive versus motor deficit contribute to the decline can help pinpoint which countermeasures to employ. A moderate or weak relationship between cognitive and fine motor test performance would indicate that each test contributes unique variance to the prediction of performance.

The first two aims of the current project were to collect normative data on both the CTB and the FMS in an astronaut-like population; specifically, certificated pilots since most U.S. and all European astronauts are pilots. These baseline, or normative, data will inform as to when an individual astronaut's performance is degraded. The third aim was to assess the correlational relationship between the two test batteries.

2. METHODS

2.1. Participants

Data were collected on n = 89 individuals between the ages of 23 to 78 (mean=37.18, SD=12.65). This total was composed of fifty military pilots (i.e., U2 Dragonlady, C5 Galaxies, KC-10 Extenders, C-17 Globemasters, F/A-18E Super Hornet Fighter), 35 private, corporate or commercial airline pilots and seven military personnel with non-flying responsibilities within military aircraft (e.g., flight engineers, gunners, loadmasters, etc.). There were seventy-eight male and 11 female participants. Seventy-four pilots were right-handed, 9 were left-handed and 6 reported being ambidextrous (although they used their right hand for data collection). Highest level of education self-reports showed 10 pilots with a high-school education, 43 with a bachelor's degree, 35 with a master's degree and one doctoral level degree. Estimated number of flight hours ranged from 55 to 21,000 (mean=2272.26, SD=3308.94) with a median of 1800 hours and mode of 3000 hours. Of the 89 participants, 87 collected data on the CTB, 88 on the FMS long version and 82 on the FMS short version. Eighty participants collected data on all three test batteries.

2.2. Equipment

CTB software was run on a Hewlett Packard ZBook with Intel(R) Core™ i7-6820HQ CPU @ 2.70GHz processor. Display size was 34.29 cm W x 17.78 cm H with a brightness setting of 100%. The display resolution was set to 1920 x 1080 viewed from a distance of approximately 18 in. System latency was defined as 42.5 ms (keyboard) and 41.2 ms (mouse) based on the average latency determined by Pulsar Informatics on similar machines. FMS software was run on a MC707LL/A Version 9.3.5 iPad with a brightness setting at approximately 75%.

2.3. Procedures

Each testing session began with an explanation of the project's overall goal. Written informed consent was obtained as per the requirements set by the NASA Ames Institutional Review Board. Each participant was assigned a random number that was associated with a randomized testing sequence of the Cognition, FMS-short or FMS-long test batteries. The CTB software provided the following fixed order of testing on the ten sub-tests.

Object Tracking (aka: Motor Praxis Test - MPT)

The Object Tracking (TRK) test determines how well participants use the computer trackpad and is a measure of visual location identification, psychomotor speed and finger dexterity. Participants are shown 20 blue squares presented one at a time at a random location on the screen. Each square is successively smaller. As soon as a square appears, participants are to use the trackpad to rapidly move the cursor onto the square and then click using the trackpad button. As soon as the participant clicks on the square it disappears and another square is presented.

Image Learning (aka: Visual Object Learning Test – VOLT)

The Image Learning (LRN) test is a test of spatial image learning and working memory retrieval of visual images. Participants are asked to remember 10 images of wireframe objects with one facet colored blue. Each image is shown successively for 5 seconds. In the test phase they are then shown 20 images, one at a time. Half of these were the learning set. Participants end the presentation by clicking on one of four options labeled "Definitely Yes", "Probably Yes", "Probably No" or "Definitely No" as to whether the image was in the learning set.

Recent Memory (aka: Fractal 2-Back - F2B)

The Recent Memory (MEM) test measures distractor effects on working memory maintenance and retrieval as well as sustained attention. Participants are shown 62 fractal patterns, one at a time, for 1.75 sec. They are to press the spacebar during the presentation if the pattern is the same as the one two patterns before. This requires remembering the last two patterns, comparing them to the current image while continuously updating their memory as the trials progress. Fifteen of the 62 test images satisfied the two-back criteria.

Concept Formation and Ranking (aka: Abstract Matching – AM)

The Concept Formation and Ranking (CON) test measures the ability to group stimuli in some meaningful way (abstraction) and to learn undisclosed, abstract concepts or rules based on feedback. Pairs of figures (circles, triangles, hexagons, crosses or stars) in one of three shades (light blue, dark blue or unfilled) are shown on the bottom of the screen. A single figure is shown in the center top of the screen. The task is to choose the figure pair at the bottom of the screen that best fits with the top figure. After clicking on their choice, the participant is provided feedback. They are to use the feedback during the 30 trials to learn the set of rules and therefore must exhibit cognitive flexibility.

Line Orientation Alignment (aka: Line Orientation Test – LOT)

The Line Orientation Alignment (LOA) test is a perceptual task. The participant is shown two lines. The reference line (6.06 cm length) is shown in random orientations and positions on the screen. The test line is shown at a randomly assigned orientation with its centroid a constant distance (6.93 cm) from the reference line. The test line may be one, of four, line lengths (1.73, 3.46, 5.19 and 6.06 cm). The task is to rotate the test line, using arrows positioned on the lower screen, in set increments of 2 deg, until it is parallel to the reference line. There are twelve, self-paced trials.

Emotion Recognition (aka: Emotion Recognition Test – ERT)

In the Emotion Recognition (EMO) test, twenty three-quarter-head shots of adults (of various ages and ethnicity) are presented one at a time. The task is to categorize their expression from a list of five emotions ("Happy", "Sad", "Angry", "Fearful" or "No Emotion").

Analogy Reasoning (aka: Matrix Reasoning Test – MRT)

The Analogy Reasoning (ANR) test measures the ability to examine an array of patterns containing one blank cell and to deduce relationships among the patterns that are satisfied by the best choice from five options to fill-in the missing cell. There are 12 trials.

Symbol Search (aka: Digit-Symbol Substitution Test – DSST)

In the Symbol Search (SYM) test participants are shown a legend of nine reference symbol-digit pairs at the bottom of the screen. A series of test symbols are presented at the top of the screen one at a time. For approximately 90 seconds, the task is to select, using the top row number keys, the number associated with the test symbol.

Risk Taking (aka: Balloon Analog Risk Test - BART)

In the Risk Taking (RSK) test participants inflate 30 balloons, shown one at a time, as much as they can without popping them. Pressing an "inflate" button increases virtual earnings by $1 or pops the balloon. The participant may press a "collect" button, rather than continuing to inflate, to transfer the current earnings into a total winnings sum. If the balloon pops, the current earnings are lost, but the accumulated winnings are untouched. A probability distribution function defines when each of the 30 balloons will pop. Participants are not informed about the probability distribution characteristics.

3-Minute Reaction Time (aka: Psychomotor Vigilance Test - PVT)

The 3-Minute Reaction Time (3RT) test measures how quickly the participant can respond to the onset of a millisecond counter. The test continues for 3-minutes. Instructions include a warning not to respond before the counter begins (false start). To aid with fixation, a continuously presented box is shown at the center of the screen. The counter displays the last reaction time for 1 sec. The next counter appears at a time sampled from a uniform distribution over 1 to 4 sec.

The FMS software, both long and short versions, also provide a fixed order of testing. Practice is provided for all sub-tests. Sub-tests and conditions included in the FMS long version follow.

Drag (Horizontal and Vertical)

The Drag test measures manual dexterity, or the speed of arm movements. The task is to push a white square back and forth or up and down from one designated area on the screen to another. Each block contained 16 trials. Each participant ran in twelve blocks of trials (2 directions x 2 repetitions x finger/stylus) for the long version plus (2 repetitions x finger/stylus) for the short version of the test.

Point (Clockwise and Counterclockwise for Short and Long Distances)

The Point test also measures manual dexterity. An annulus of squares is presented. The task is to tap the highlighted square. The top square is always highlighted first. An arrow indicates if the highlighted square will travel clockwise or counterclockwise around the annulus. As soon as the participant taps the highlighted square, it is de-emphasized and the square on the opposite side of the annulus is highlighted. Each block contained 18 trials. Each participant ran in twenty blocks of trials (2 directions x 2 repetitions x 2 annulus sizes x finger/stylus) for the long version plus (2 repetitions x finger/stylus) for the short version of the test.

Trace (Clockwise and Counterclockwise for a Circle and a Square)

The Trace test also measures manual dexterity. The participant follows the outline of a geometric figure starting at the location of a small circle labelled "Start". They trace along the outline in the direction indicated by an arrow. Feedback was provided on the path traced. Each block contained 5 trials. Each participant ran in twelve blocks of trials (2 directions x 2 repetitions x finger/stylus) for the long version plus (2 repetitions x finger/stylus) for the short version of the test.

Pinch-Rotate (0 and 45-deg Rotation)

The Pinch-Rotate task measures finger dexterity and wrist-finger speed. In this test the participant places their pointer finger and thumb on two circles at the opposite corners of a blue square. They then pinch and rotate the square on the iPad screen to align with a 45-deg rotated inner black square. When the two squares are coincident, the participant lifts their fingers. The adjustable blue square location was provided as it is rotated. Each block contained 6 trials. Each participant ran in five blocks of trials (2 orientations x 2 repetitions) for the long version plus 1 repetition for the short version of the test.

Data were obtained with both the participant's pointer finger and a stylus for all but the Pinch-Rotate test. The FMS long version included a repetition of all tests. The FMS short version excluded the long-distance Pointing test, the square Tracing test and the 45-deg Pinch-Rotate test.

3. RESULTS

3.1. Normative Data on the CTB

Table 1 shows means and standard deviations for accuracy scores and response times for each test in the Cognitive Test Battery. Acronyms used to refer to these tests in previous publications are also provided.

Table 1. Cognition Test Battery Means and Standard Deviations (in parentheses) for 89 Participants.

CTB Sub-test	Former Acronym	Accuracy Score	Response Time, sec
Object Tracking TRK	MPT	-0.5220 (0.1032)	1.2153 (0.2034)
Image Learning LRN	VOLT	0.8650 (0.1052)	2.599 (0.593)
Recent Memory MEM	NBACK, F2B	0.7683 (0.1070)	0.606 (0.093)
Concept Formation CON	AM	0.669 (0.162)	2.767 (1.216)
Line Orientation Alignment LOA	LOT	0.1134 (0.0290)	0.8453 (0.2694)
Emotion Recognition EMO	ERT	0.633 (0.118)	3.523 (1.567)
Analogy Reasoning ANR	MRT	0.684 (0.131)	11.538 (3.283)
Symbol Search SYM	DSST	0.986 (0.025)	1.468 (0.288)
Risk Taking RSK	BART	2.4805 (0.1683)	3.070 (1.231)
Reaction Time 3RT	PVT	0.949 (0.043)	0.2107 (0.0168)

The accuracy scores for Concept Formation, Emotion Recognition, Analogy Reasoning, and Symbol Search are the proportion of correct responses. The Image Learning score is the area under the Receiver-Operating-Characteristic (ROC) curve based on the three Hit and False Alarm rates from the four rating categories. The Recent Memory accuracy score is the average of the proportions correct on “same-2-back” and “not-same” trials. The Reaction Time score is the number of correct responses over the number of trials plus the number of anticipation responses. The Line Orientation Alignment accuracy score is an estimated slope of the psychometric function in proportion/degrees. The Object Tracking accuracy score is the average distance from the first response to the box center, divided by the half of the box width. Finally, Risk Taking performance is the expected winnings per trial based on an estimated distribution of how many times they pumped before collecting.

For Concept Formation, Analogy Reasoning, Image Learning, and Emotion Recognition, the response time is simply the average time to the response on each trial. For Object Tracking it is the time to the first response. The Recent Memory response time is the average time to “same-2-back” responses. The Line Orientation Alignment response time is the average response time for individual stimuli with zero rotation difference or absolute rotation differences that were less than 12 deg and that led to at least one error. The Risk Taking response time is the average time to the response that ended the trial. Finally, Reaction Time response time is the average of times greater than 0.1 sec (not anticipations) and less than 0.355 sec (not lapses).

Correlations among the accuracy scores and correlations among the response times are shown separately in Table 2, with the accuracy correlations below the main diagonal and the response time correlations above. Correlations significantly different from zero at the 1% level (two-tailed) are shown with double asterisks. Single asterisks are provided for significant correlations at the 5% level.

Only 2 of the 45 accuracy score correlations are significant at the 1% level, while 25 of the response time correlations are significant at that level, indicating a much higher level of commonality among the response times.

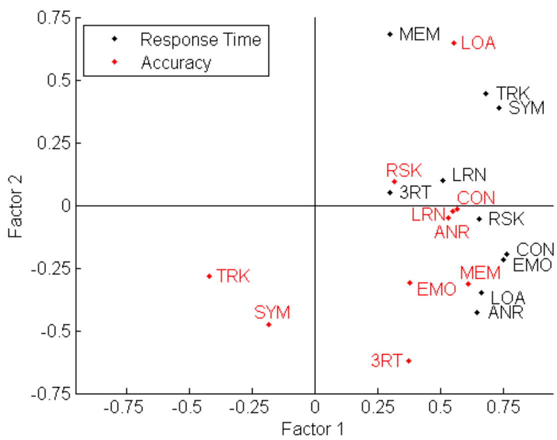

Figure 1. Factor Loadings on the first two factors for both speed and accuracy correlations on the CTB

Principal component factor analyses were performed on both the speed and accuracy correlations. Factor loadings on the first two factors are shown in Figure 1. The proportion of variance accounted for by Factors 1 and 2 on response time were 0.4031 and 0.1195, and on accuracy were 0.2052 and 0.1304, respectively. The high proportion on the first response time factor corresponds to the higher correlations among the response times. All of the response time measures have positive weighting on their first principal component. The factor loadings on the first-time factor for the Reaction Time and Recent Memory time scores are much lower than those of the other eight time scores. The accuracy score first principal component mainly separates perceptual-motor tasks (LOA, TRK, SYM) from memory and reasoning tasks (MEM, CON, LRN). The SYM and PVT accuracy scores have such high mean proportions (0.986 and 0.949) that they are unlikely to be very reliable. The emotion recognition test (EMO) groups with the more cognitive tests.

Table 2. Correlations among response times (above main diagonal) and accuracy (below main diagonal) for each CTB sub-test.

	TRK	LRN	MEM	CON	LOA	EMO	ANR	SYM	RSK	3RT
TRK	----	0.34**	0.25*	0.39**	0.28**	0.34**	0.21*	0.63**	0.50**	0.2
LRN	0.04	----	0.1	0.26*	0.28**	0.40**	0.17	0.33**	0.27*	-0.01
MEM	-0.23*	0.07	----	0.15	-0.02	0.17	0.11	0.35**	0.09	0.08
CON	-0.1	0.18	0.23*	----	0.40**	0.57**	0.58**	0.48**	0.43**	0.19
LOA	0.02	0.2	0.16	0.24*	----	0.46**	0.43**	0.42**	0.48**	0.13
EMO	-0.05	0.06	0.16	0.13	0.02	----	0.56**	0.43**	0.32**	0.17
ANR	-0.1	0.05	0.31**	0.20	0.30**	0.16	----	0.30**	0.30**	0.12
SYM	-0.13	-0.01	-0.05	-0.07	0.20	-0.02	0.11	----	0.30**	0.15
RSK	0.05	0.23*	-0.01	-0.03	0.05	0.01	0.06	0.20	----	0.25*
3RT	0.01	0.21	0.30*	0.12	0.09	0.18	0.03	-0.03	0.25*	----

3.2. Normative Data on the FMS

Table 3. FMS Means of the Median Latencies (sec) and Standard Deviations (in parentheses) for 82 Participants.

# Blocks	Drag *12*	Point *20*	Trace *12*	Rotate *5*
Average Latencies (sec)	0.6317 (0.1037)	0.4885 (0.0607)	3.4903 (1.6575)	2.2485 (0.5531)

Table 3 presents the average latencies for each FMS sub-test. To compute these latencies, the median score for each of 82 participants was computed for each block. The mean of these medians was then computed for each participant. Finally, the mean was computed over the participants.

Table 4. Correlations among Response Times and Factor Analysis Factor Loadings for each FMS sub-test.

	Drag	Point	Trace	Rotate
Drag	1.0000	0.6386	0.2090	0.3727
Point	0.6386	1.0000	0.0378	0.2853
Trace	0.2090	0.0378	1.0000	0.4801
Rotate	0.3727	0.2853	0.4801	1.0000
Factor 1 Loading	0.8209	0.7278	0.5343	0.7414
Factor 2 Loading	-0.3533	-0.5596	0.7299	0.4145

Correlations among the latencies are shown in Table 4. To summarize these results, principal component factor analyses were performed on the correlations. The factor loadings on the first two factors are also provided. The proportion of variance accounted for by Factors 1 and 2 are 0.5097 and 0.2856, respectively. The first principal component reflects a strong correlation among the tests. The second factor separates the Drag and Point tasks from the Trace and Rotate tasks.

3.3. Motor Contributions for each CTB Sub-Test

Table 5. Correlations between FMS Factor 1 and the CTB Response Times & Accuracy for each CTB sub-task

Sub-task	Response Time	Accuracy
TRK	0.3198**	0.0504
LRN	0.1190	-0.2980**
MEM	0.0932	-0.1461
CON	0.3662**	-0.0558
LOA	0.3301**	-0.0528
EMO	0.2581*	-0.0749
ANR	0.4436**	0.0648
SYM	0.3082**	0.2552*
3RT	0.1216	0.0110

To evaluate the relationship between the CTB and the FMS, the correlations were computed between FMS Factor 1 (loading coefficients in Table 4) and the response time and accuracy scores from each CTB sub-task (Table 5). Correlations are given asterisks in Table 5 as in Table 2. The response time correlations are similar to the CTB response time factor loadings in Figure 1. The Reaction Time time has a small correlation, as do the Image Learning and Recent Memory response times. The only accuracy score correlation to reach the 0.01 level of significance is the negative correlation of the Image Learning accuracy with FMS Factor 1. The second highest correlation is the positive correlation of Symbol Search accuracy with FMS Factor 1.

4. DISCUSSION

Data were collected in 89 armed forces, commercial and private pilots on two NASA-developed test batteries. These data characterize what is usual in an astronaut-like population, permitting the identification of potential decline in crew performance during spaceflight.

Our results are similar to those of Moore et al. (2017).[14] They collected data on the CTB in 96 Philadelphians (age range from 25 to 56 years) with at least a Masters' degree in science, technology, engineering or mathematics. The current CTB normative data are consistent with Moore et al. (2017) for all but the ERT and the BART. This deviation is likely due to the differences in CTB versions used in their study as compared to ours.

High correlations among measures of CTB speed were also reported by Gur et al. (2010)[15]. We found the ability to perform fine motor behaviors rapidly makes a significant contribution toward the ability to perform the CTB rapidly. This ability cannot be simple motor speed because the reaction time does not correlate with the ability to perform the other tasks rapidly. It only received a small loading on the first factor extracted from the cognitive test score speeds and it has a small correlation with the first FMS time factor.

ACKNOWLEDGEMENT

Significant contributions were received from many individuals, staff and service members at Beale Air Force Base, Lemoore Naval Air Station, Travis Air Force Base, Edwards Air Force Base, West Valley Flying Club and pilot bloggers. This directed project was funded by the NASA Human Research Program, Human Factors and Behavioral Performance Element.

5. REFERENCES

1. Basner, Mathias, Adam Savitt, Tyler M. Moore, Allison M. Port, Sarah McGuire, Adrian J. Ecker, Jad Nasrini et al. "Development and validation of the cognition test battery for spaceflight." *Aerospace medicine and human performance* 86, no. 11 (2015): 942-952.

2. Straker, Leon, and Svend Erik Mathiassen. "Increased physical work loads in modern work–a necessity for better health and performance." *Ergonomics* 52, no. 10 (2009): 1215-1225.

3. Perry, Carlene M., Mohamed A. Sheik-Nainar, Noa Segall, Ruiqi Ma, and David B. Kaber. "Effects of physical workload on cognitive task performance and situation awareness." *Theoretical Issues in Ergonomics Science* 9, no. 2 (2008): 95-113.

4. Sanes, Jerome N., and Edward V. Evarts. "Effects of perturbations on accuracy of arm movements." *Journal of Neuroscience* 3, no. 5 (1983): 977-986.

5. Conway, G. E., J. L. Szalma, and P. A. Hancock. "A quantitative meta-analytic examination of whole-body vibration effects on human performance." *Ergonomics* 50, no. 2 (2007): 228-245.

6. Ross, Helen E. "Motor skills under varied gravitoinertial force in parabolic flight." *Acta Astronautica* 23 (1991): 85-95.

7. Grabherr, Luzia, and Fred W. Mast. "Effects of microgravity on cognition: The case of mental

imagery." *Journal of Vestibular Research* 20, no. 1, 2 (2010): 53-60.

8. Barnard, Kirsten E., Joshua J. Broman-Fulks, Kurt D. Michael, Rosemary M. Webb, and Laci L. Zawilinski. "The effects of physiological arousal on cognitive and psychomotor performance among individuals with high and low anxiety sensitivity." *Anxiety, Stress, & Coping* 24, no. 2 (2011): 201-216.

9. Barnard, Kirsten E., Joshua J. Broman-Fulks, Kurt D. Michael, Rosemary M. Webb, and Laci L. Zawilinski. "The effects of physiological arousal on cognitive and psychomotor performance among individuals with high and low anxiety sensitivity." *Anxiety, Stress, & Coping* 24, no. 2 (2011): 201-216.

10. Flindall, Ian Richard. "Acute mental fatigue and cognitive performance in the medical profession." Thesis submitted to the Imperial College of London. (2015).

11. Satish, Usha, Mark J. Mendell, Krishnamurthy Shekhar, Toshifumi Hotchi, Douglas Sullivan, Siegfried Streufert, and William J. Fisk. "Is CO2 an indoor pollutant? Direct effects of low-to-moderate CO2 concentrations on human decision-making performance." *Environmental Health Perspectives* 120, no. 12 (2012): 1671-1677.

12. Williamson, Ann M., and Anne-Marie Feyer. "Moderate sleep deprivation produces impairments in cognitive and motor performance equivalent to legally prescribed levels of alcohol intoxication." *Occupational and Environmental Medicine* 57, no. 10 (2000): 649-655.

13. Newman, Dava J., and Corinna E. Lathan. "Memory processes and motor control in extreme environments." *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 29, no. 3 (1999): 387-394.

14. Moore, Tyler M., Mathias Basner, Jad Nasrini, Emanuel Hermosillo, Sushila Kabadi, David R. Roalf, Sarah McGuire et al. "Validation of the cognition test battery for spaceflight in a sample of highly educated adults." *Aerospace medicine and human performance* 88, no. 10 (2017): 937-946.

15. Gur, Ruben C., Jan Richard, Paul Hughett, Monica E. Calkins, Larry Macy, Warren B. Bilker, Colleen Brensinger, and Raquel E. Gur. "A cognitive neuroscience-based computerized battery for efficient measurement of individual differences: standardization and initial construct validation." *Journal of Neuroscience Methods* 187, no. 2 (2010): 254-262.