MILITARY MEDICINE, 178, 4:365, 2013

Defense Automated Neurobehavioral Assessment (DANA)—Psychometric Properties of a New Field-Deployable Neurocognitive Assessment Tool

Corinna Lathan, PhD; James L. Spira, PhD, MPH†; Joseph Bleiberg, PhD‡; Jack Vice*; CAPT Jack W. Tsao, USN§*

ABSTRACT The Defense Automated Neurobehavioral Assessment (DANA) is a new neurocognitive assessment tool that includes a library of standardized cognitive and psychological assessments, with three versions that range from a brief 5-minute screen to a 45-minute complete assessment. DANA is written using the Android open-source operating system and is suitable for multiple mobile platforms. This article presents testing of DANA by 224 active duty U.S. service members in five operationally relevant environments (desert, jungle, mountain, arctic, and shipboard). DANA was found to be a reliable instrument and compared favorably to other computer-based neurocognitive assessments. Implications for using DANA in far-forward military settings are discussed.

INTRODUCTION

In January 2009, the U.S. Navy Bureau of Medicine and Surgery identified a need to enhance existing battlefield concussion assessment and requested the development of a durable, portable, and field-hardened neurocognitive assessment tool (NCAT) to provide a practical means to conduct neurocognitive and psychological assessment in field deployment settings. The purpose of combining neurocognitive and psychological assessment was to permit more comprehensive evaluation of the broad range of problems that may be encountered during combat deployment. This article describes the resulting NCAT, Defense Automated Neurobehavioral Assessment (DANA); DANA's psychometric properties based on assessment of 224 active duty U.S. service members under challenging field conditions; and presents comparisons to published NCAT data.

DANA consists of three test batteries of different durations and compositions designed for increasingly detailed assessment (Table I). The three batteries include (1) DANA Rapid, a 5-minute battery of three basic reaction-time measures; (2) DANA Brief, a 15-minute test that includes DANA Rapid plus additional neurocognitive tests as well as psychological screening tools for post-traumatic stress disorder (PTSD), depression, and insomnia; and (3) DANA Standard, a 45-minute more comprehensive battery of neurocognitive and psychological tests. This hierarchical set of batteries is designed to facilitate health care providers' access to standardized, reliable, and valid objective and subjective measures. DANA's portability, multiple test batteries, and user-friendly interface enable its use by a wide range of health care providers, from frontline medics/corpsmen to licensed health care professionals.

Establishing reliability and feasibility of this platform in a military population is necessary before attempting clinical validation and utilization. The eventual goal of DANA is to assist clinicians to (a) make rapid and accurate assessment of cognitive and psychological dysfunction secondary to brain injury and/or the psychological wounds and stressors of war,

*AnthroTronix, 8737 Colesville Road, Suite L-203, Silver Spring, MD 20910.

†Department of Veterans Affairs, National Center for PTSD, Pacific Islands Health Care System, 3375 Koapaka Street, Honolulu, HI 96819.

‡National Intrepid Center of Excellence, Walter Reed National Military Medical Center, 8901 Wisconsin Avenue, Building 51, Bethesda, MD 20889.

§Wounded, Ill and Injured Directorate, US Navy Bureau of Medicine and Surgery, 7700 Arlington Boulevard, Falls Church, VA 22032.

The preliminary data of this article, "Defense Automated Neurobehavioral Assessment (DANA): A Field-Deployable Tool for Assessing Concussion and Deployment Stress," was presented as an oral presentation by the authors J.L.S., C.L., J.B., and J.W.T. at the Military Health System Research Symposium (MHSRS) Conference, Fort Lauderdale, FL, August 2012.

doi: 10.7205/MILMED-D-12-00438

TABLE I. DANA Test Batteries*

DANA Rapid (5 Minutes)	DANA Brief (15 Minutes)	DANA Standard (45 Minutes)**
Simple Reaction Time (SRT)	SRT	SRT
Procedural Reaction Time (PRO)	Code Substitution Simultaneous (CDS)	CDS
Go/No-Go (GNG)	PRO	PRO
	Spatial Discrimination (SPD)	SPD
	GNG	GNG
	Code Substitution Delayed (CDD)	CDD
	SRT	Matching to Sample (MSP)
	Patient Health Questionnaire (PHQ)	Sternberg Memory Search (STN)
	Primary Care PTSD Screen (PC-PTSD)	SRT
	Insomnia Screening Index (ISI)	Combat Exposure Scale (CES)
		PHQ
		Pittsburgh Sleep Quality Index (PSQI)
		PTSD Checklist—Military Version (PCL-M)
		Deployment Stress Inventory (DSI)

*For detailed test descriptions, see Table AI. **MSP and STN were still under development at the time of this testing and so are not included in the results.

(b) facilitate referral to treatment for wounded service members, (c) monitor recovery, and (d) aid in return-to-duty determination. Thus, DANA is intended to enhance military capability and better ensure a healthy fighting force.

METHODS

DANA Platform

DANA is a Java-based mobile application that runs on an Android operating system. The primary advantages of Android are that it is open source, open license (Apache 2.0), well supported and based on a Linux kernel, which is nearly ubiquitous. Java has the advantage of being a high-level, class-based, object-oriented language designed as a "write once, run anywhere" solution and thus is portable across a wide range of devices and desktops. DANA, therefore, can run on any Android mobile device and can be used with a stylus or touch screen.

Based on the Navy Bureau of Medicine and Surgery, requirements for MIL-SPEC commercial-off-the-shelf hardware, we conducted a comprehensive trade study and selected the Trimble Nomad, the military-grade-hardened handheld computer used in the current study. A Tektronix 100 MHz analog to digital (ADC) oscilloscope was used to test the input variability of device hardware and device software that could contribute to the overall response times. A push action switch was connected to the ADC, which was then used as the input stylus on the Nomad to measure RT. The interval between two inputs as recorded by DANA and by the ADC was compared over 10 trials. The average difference was 6.8 milliseconds with a standard deviation (SD) of 3.7 milliseconds. By comparison, the input variability with a Microsoft windows personal computer can range from 4–25 milliseconds.[1,2]

DANA Test Battery

Selection of the neurocognitive and psychological tests included in DANA was established by a tri-service, Veterans Administration, and civilian scientific advisory board that included military and civilian neuropsychologists and psychologists, neurologists, and corpsmen. All tests included in DANA's test batteries meet the requirements of the American Psychological Association's standard for tests and measurements and all tests are in the public domain. Eight cognitive tests and seven psychological questionnaires were selected (Table AI) and are divided into three test batteries, as shown in Table I. Tests were selected based upon their potential to address specific deployment-related concerns, such as concussion and combat distress or exhaustion. Although all tests utilized have an extensive literature regarding their reliability and validity, they have not been reported in this specific configuration nor implemented for service members in this manner. The advisory board also contributed to and provided feedback on the user interface design to ensure ease of use by multiple levels of caregivers including the corpsmen, general medical officers, and neuropsychologists.

Participants and Procedure

To evaluate the deployment feasibility of DANA, we recruited 224 active duty service members comprising 40 or more active duty military personnel in each of 5 diverse operational environments. No subjects were excluded, since all service members were fit for duty, not undergoing any disability evaluation, and thus assumed to be healthy. The purpose of assessing service members across diverse environmental conditions was to show the robustness of the hardware and software administration under different operational tempos and to identify any environmental concerns with the reliability of the instrument.

- — Arctic (Thule Air Force Base—Greenland in the winter)
- — Jungle (U.S. Marine Corps Jungle Warfare Training Center—Okinawa, Japan, in the summer)
- — Altitude (U.S. Marine Corps Mountain Warfare Training Center—Bridgeport, CA, approximately 3,000 m)
- — Desert (U.S. Marine Corps Desert Warfare Training Center—Twentynine Palms, CA, in the summer)
- — Shipboard (USS George Washington during high seas in the Western Pacific)

Device performance (e.g., battery life, display characteristics) was evaluated under the specific environmental conditions (e.g., humidity, temperature) through a minimum of 12-hour exposure. The only instrumentation reliability issue was a screen refresh rate delay in the Go/No-Go (GNG) test in the arctic environment. Because this screen rendering delay would affect test results, the rendering process software was redesigned, which successfully mitigated the delay.

The research protocol was approved by the AnthroTronix Institutional Review Board, the VA Institutional Review Board, and received a Department of the Navy Human Research Protections Program review. A letter was obtained from the commanding officer of each test facility and all subjects signed an informed consent document to participate in testing. On Day 1, each subject was tested on all three batteries, the DANA Rapid, Brief, and Standard. Subjects returned on Day 2 to repeat the sequence of batteries resulting in the following protocol:

— Day 1 (approximately 120 minutes)—Consent Process, DANA Rapid, DANA Brief, DANA Standard
— Day 2 (approximately 40 minutes)—DANA Rapid, DANA Brief, DANA Standard (cognitive tests only)

(The above times include 5-minute breaks between each battery.)

A research team of clinical psychologists and technical staff administered testing. Participants were instructed to hold the stylus about 1-cm above the screen, and to respond as rapidly and accurately as possible. All other instructions were embedded within the tests. To minimize learning and practice effects, test stimuli are generated at random and each test has practice trials before the actual test trials. Parameters of the final version of DANA's individual subtests are described in Table AII including each subtest's stimulus presentation duration, response interval, and interstimulus interval.

Data were analyzed in SPSSv20 for descriptive statistics, split-half reliability, test–retest reliability, and cross-test correlations. For internal consistency, we examined split-half correlations of the first and second half set of trials for the first administration of each test on each day. To evaluate test–retest reliability across administrations, we calculated intraclass correlation coefficients (ICCs)[3] that have been used to evaluate reliability for other health status instruments.[4] Because of multiple analyses, significance levels were set to between $p < 0.01$ and $p < 0.001$, depending upon the number

TABLE II. Descriptive Statistics for all DANA Variables for Each Administration

	Task	Administration	*n*	Median RT Correct ± SD		Average of Median Throughput ± SD		Percentage Correct ± SD	
1	SRT	1	223	309.7	±65.3	199.6	±33.4	99.7	±3.3
		2	223	309.3	±64.6	199.8	±34.6	99.5	±2.6
		3	220	300.6	±55.5	204.4	±33.8	99.4	±3.4
		4	213	302.0	±53.3	202.2	±33.4	99.3	±4.0
		5	212	308.3	±65.1	200.6	±36.7	99.3	±4.7
		6	172	298.4	±68.5	207.1	±31.7	99.7	±2.4
		7	172	307.8	±77.0	202.1	±35.7	99.4	±2.3
		8	164	305.1	±49.2	200.7	±30.6	99.6	±2.2
		9	122	317.9	±69.0	195.7	±37.5	99.7	±1.5
		10	121	310.7	±54.9	197.8	±30.7	99.8	±0.8
2	PRO	1	224	604.5	±101.6	100.1	±15.5	98.3	±4.3
		2	220	579.6	±91.4	103.8	±15.1	98.1	±4.0
		3	213	571.8	±84.8	104.8	±14.2	98.0	±4.0
		4	174	579.1	±79.9	103.6	±13.4	98.3	±3.1
		5	164	579.7	±95.4	104.1	±15.6	98.2	±3.5
		6	122	565.0	±84.1	105.7	±15.0	97.6	±5.3
3	GNG	1	214	535.4	±96.8	114.2	±18.5	99.1	±2.5
		2	214	519.3	±86.8	117.6	±18.5	99.3	±2.3
		3	193	520.0	±91.4	116.7	±19.9	98.2	±4.5
		4	163	527.2	±97.4	116.1	±19.2	99.0	±2.8
		5	99	506.4	±98.3	120.9	±20.8	98.8	±3.6
		6	75	521.0	±109.6	117.7	±21.7	98.4	±5.2
4	SPD	1	221	1690.2	±376.3	34.4	±7.1	92.8	±5.4
		2	209	1533.0	±361.5	37.3	±9.4	90.6	±5.7
		3	172	1562.3	±383.9	37.8	±9.6	93.1	±5.8
		4	119	1456.6	±311.4	38.2	±7.4	89.3	±6.2
5	CDS	1	223	1284.9	±277.7	47.5	±9.5	97.5	±3.1
		2	212	1256.7	±234.5	47.8	±8.9	96.8	±4.0
		3	171	1228.7	±257.0	49.8	±10.8	97.7	±3.1
		4	120	1193.4	±211.1	50.5	±9.4	97.3	±3.2
6	CDD	1	212	1046.7	±221.5	55.0	±11.9	92.1	±8.2
		2	190	1004.3	±183.1	56.2	±11.7	91.1	±9.1
		3	161	996.7	±184.2	56.8	±12.8	90.9	±9.5
		4	103	956.9	±157.4	59.6	±11.5	92.5	±7.9

TABLE III. Comparison to Previously Reported Data[a]

		DANA	ANAM (2006)	ANAM (2008)	ANAM (2012)
SRT	*n*	223	2,261	5,237	107,413
	Mean[a]	309.7	261.3	267	261
	SD	65.3	56.1	74	47
	Ratio	21.08%	21.47%	27.72%	18.01%
PRO	*n*	224	—	—	107,353
	Mean[a]	604.5	—	—	592
	SD	101.6	—	—	90
	Ratio	16.81%	—	—	15.20%
CDS	*n*	223	2,331	5,237	107,546
	Mean[a]	1284.9	1,191	1,096	1,162
	SD	277.7	248.7	265	272
	Ratio	21.61%	20.88%	24.18%	23.41%
CDD	*n*	212	1,891	5,202	107,523
	% Accuracy	92.1	88.7	86.30	90
	SD	8.2	9.3	15.80	11.4
	Ratio	8.90%	10.48%	18.31%	12.67%

[a]Medians are shown for DANA.

of analyses conducted to correct for type-1 error. Psychological measures were scored using conventional methods.

RESULTS

DANA performed well in all five field environments with no significant difference across data sets; therefore, the data for all five operational environments were combined. The number of total subjects for each test ranged from $n = 75$ to $n = 224$ depending on whether or not the service member was available to participate in all administrations across 2 days. All but one test (Code Substitution Delayed [CDD]) had over 200 subjects for at least two test administrations.

Scores on psychological measures revealed an overall psychologically healthy sample. Combat Exposure Scale (CES) was in the light range (3.7, with 17 indicating moderate exposure). PTSD Checklist—Military Version (PCL-m) (26), Patient Health Questionnaire (PHQ)-8 (4.6), Pittsburgh Sleep Quality Index (PSQI) (5.9), and Deployment Stress Inventory (DSI) (9.7) were each far below the score needed to reach clinical criteria.

Data exclusions included the elimination of trials with anticipatory responses and test administrations indicating random responses (criterion for exclusion was set at less than 65% correct—since tasks are binary, 50% correct represents random responding). No slow RT responses were eliminated, but to mitigate their undue influence we used medians rather than means to describe the data. Less than 1% of the response trials and less than 2% of the test administrations were eliminated based on these criteria. No subjects were eliminated from analysis based upon a criterion of having more than one test administration eliminated in a battery.

Table II shows the descriptive data for each administration of each test, and Table III compares DANA median RT and SD to previously published reports of mean data from the Automated Neuropsychological Assessment Metrics (ANAM, currently used by the U.S. Department of Defense for baseline, predeployment neurocognitive testing) in military personnel.[5–7] DANA uses median correct RT as the relevant metric, whereas ANAM reports means; however, because of the sample size and the normal populations assessed, the mean and median are assumed to be similar for ANAM.

Although not significantly different, absolute RT values may be different because of differing instrumentation (stylus versus mouse button) between ANAM and DANA. It is also possible to use the published ANAM data to calculate Coefficients of Variation (CVs) (ratio of RT to SD of RT) to show stability[8] and these are compared to analogous DANA subtests in Table III.

For reliability within administrations, split-half correlations of the odd–even trials are reported for the first administration of day 1 and day 2 for each test. Correlations were acceptably high ($p < 0.001$) and generally above 0.85—Simple Reaction Time (SRT) (0.91, 0.93), Procedural Reaction Time (PRO) (0.87, 0.86), GNG (0.85, 0.85), Code Substitution Simultaneous (CDS) (0.94, 0.93)—except for Spatial Discrimination (SPD) (0.76, 0.76) and CDD (0.76, 0.82). For test reliability across administrations, the ICC was calculated. The ICC for SRT was 0.95 indicating excellent reliability across the 10 administrations over 2 days. For PRO, GNG, SPD, and CDS, similar reliability was achieved with ICC values of 0.91, 0.95, 0.89, and 0.88, respectively. Only the CDD test did not have high reliability (0.54), which is expected for repeat CDD tests within a short time period because of the change in codes with each administration, resulting in proactive memory interference.

Correlations were also conducted between all psychological and cognitive measures. With the exception of CDD (difficulties using multiple administrations of this subtest because of memory interference were noted above), all cognitive subtests ($p < 0.001$) and psychological tests ($p < 0.001$) were correlated with each other; however, the psychological and cognitive tests did not correlate with each other in this

TABLE AI. DANA Test Descriptions

Test Name (Abbreviation)	Task Structure	Task Purpose
Simple Reaction Time (SRT)[a–c]	The subject taps on the location of the yellow asterisk symbol as quickly as possible each time it appears.	This task measures pure RT.
Procedural Reaction Time (PRO)[a–c]	The screen displays one of four numbers for 3 seconds. The subject presses on a left button ("2" or "3") or right button ("3" or "4") depending upon the number pressed.	A choice RT measure of accuracy, RT, and impulsivity. This choice RT task targets simple executive functioning with easy decision-making capabilities.
Go/No-Go (GNG)[a–c]	This is a forced choice RT task relevant to warfighters. A house is presented on the screen with several windows. Either a "friend" (green) or "foe" (white) appears in a window. The respondent must push a "fire" button only when a "foe" appears.	A choice RT measure of sustained attention and impulsivity. The test assesses speed and accuracy of targets, omissions, and commissions.
Spatial Discrimination (SPD)[b,c]	Pairs of four-bar histograms are displayed on the screen simultaneously, and the subject is requested to determine whether they are identical. One histogram is always rotated either ±90 degrees with respect to the other histogram.	Assesses visuospatial analytic ability.
Code Substitution Simultaneous (CDS)[b,c]	Subjects refer to a code set of 9 symbol-digit pairs that are shown across the upper portion of the screen. A sequence of single symbol-digit pairs is shown below the key, and the subject indicates whether or not the single pair matches the code by pressing Yes or No.	Assesses visual scanning and attention, learning, and immediate recall.
Code Substitution Delayed (CDD)[b,c]	After a delay of several intervening tests, the same symbol-digit pairs are presented without the code. The subject indicates whether or not the pairing was included in the code that was presented in the earlier code substitution learning section.	Assesses learning and short-term memory.
Sternberg Memory Search (STN)[c]	The subject memorizes a set of five letters, after which letters appear on the screen one at a time, and the subject determines if the letter on the screen is a member of the memory set.	Assesses working memory.
Matching to Sample (MSP)[c]	A single 4 × 4 checkerboard pattern is presented on the screen for brief study period. It then disappears for 5 seconds, after which two patterns are presented side-by-side. The subject indicates which of these two patterns matches the first.	A measure of short-term memory, attention, and visuospatial discrimination.
Insomnia Screening Index (ISI)[b]	A 5-item scale evaluating perceived insomnia severity and sleep habits. Each item is rated on a 5-point scale (0–4).	The total score ranges from 0 to 28 and higher scores indicate more severe insomnia. A cutoff score of 10 has been shown to indicate insomnia.[11]
Primary Care PTSD Screen (PC-PTSD)[b]	Four screening questions designed for use in clinical settings to screen for PTSD, with 3 out of 4 endorsed items suggestive of likely PTSD.	Questions assess hyper-arousal, re-experiencing, and avoidance for PTSD screening. This test is more sensitive than specific, but correlates highly with the PCL.[12,13]
Patient Health Questionnaire (PHQ)[b,c]	A 9-item depression scale assessing symptom severity and diagnostic criteria for major depressive disorder. For research purposes, item no. 9 (concerning suicide) was not included, yet research indicates that the scoring, reliability, and clinical validity are almost identical.	A score of 0–9 is likely to have no depression, 10–14 mild depression, 15–19 moderate depression, and 20+ severe depression.[14]
Pittsburgh Sleep Quality Index (PSQI)[c]	19 self-rated items and 5 partner-rated items, which measure sleep quality during the previous month. This scale differentiates "good" from "poor" sleepers based on seven areas: subjective sleep quality, sleep latency, sleep duration, habitual sleep efficiency, sleep disturbances, use of sleeping medication, and daytime dysfunction over the last month.	This scale is the most widely utilized sensitive and specific self-report measure for insomnia. A score above 6 indicates a "poor" sleeper, and a score above 12 is associated with "insomnia."[15]

(continued)

TABLE AI. Continued

Test Name (Abbreviation)	Task Structure	Task Purpose
Combat Exposure Scale (CES)[c]	A 7-item self-report measure that assesses wartime stressors experienced by service members. The total CES score (ranging from 0 to 41) is calculated by using a sum of weighted scores, which can be classified into 1 of 5 categories of combat exposure ranging from "light" to "heavy."	This scale rates cumulative combat exposure and is highly predictive of PTSD, pain and injury, TBI, depression, and other behavioral sequellae.[16]
PTSD Checklist—Military Version (PCL-M)[c]	A 17-item scale assessing symptoms in response to stressful military experiences. This scale assesses PTSD, with subscales including re-experiencing, avoidance/numbing, and hyperarousal.	Higher scores indicate increased PTSD symptomatology. In a military population, scores >49 are likely to have PTSD. For greatest specifically, scores >44 with 3 re-experiencing, 1 avoidance/numbing, and 2 Hyperarousal endorsed as at least "most of the time" are more specific for PTSD and correlate very highly (0.92) with the Clinician Administered PTSD Scale (CAPS).[17]
Deployment Stress Inventory (DSI)[c]	Based upon the neurobehavioral symptom inventory with additional items added to assess anger, pain, and sleepiness. Test is a 28-item experimental scale that factors into three domains (cognitive-emotional, somatic, and anger) and five subscales (cognitive, emotional, pain, sleep, and anger).	This experimental measure is intended to be used as a broad psychological screening tool sensitive to combat-related distress, especially reporting of persistent postconcussive symptoms.[18]

[a]DANA Rapid Battery. [b]DANA Brief Battery. [c]DANA Standard Battery.

sample of nonimpaired service members. The CES correlated mildly with the PCL-m and DSI ($p < 0.01$).

DISCUSSION

The data presented here represent a first empirical examination of the DANA tool, a portable NCAT that includes both cognitive and psychological tests. Feasibility, reliability, and internal validity were assessed through the administration of DANA to 224 active duty service members (officers and enlisted) from the U.S. Air Force, Navy, and Marine Corps across five extreme environments.

As can be seen in Table II, scores were stable across administrations. Split-half reliability correlations for DANA subtests are within acceptable ranges, and are comparable to those reported for other similar NCAT subtests.[9] ICC measures across administrations were excellent and generally exceeded those reported in similar NCAT subtests although they were comparable to previously reported aggregate NCAT scores.[10] We are currently developing more sophisticated scoring and statistical approaches to score each subtest as well as to aggregate subtests into composite index scores.

Descriptive analysis of the psychological subtests shows that the sample scored, in general, well below scores indicative of clinical problems. From the range of CES scores, it appears that most of the sample had not been exposed to moderate or greater combat. With regard to mean psychological scores and the relevant scoring for cognitive variables, all psychological variables correlate highly with each other, with the exception of the Insomnia Severity Index (ISI) and Primary Care (PC)-PTSD, which had moderate correlations, likely due to having a low number of items (4 and 7 items,

TABLE AII. Test Parameters

	No. of Response Trials	Stimulus Presentation Time (milliseconds)	Maximum Response Time (milliseconds)	Intertrial Interval (milliseconds)
SRT	40	900	900	600 to 3,000
PRO	32	2,000	2,000	500 to 1,000
GNG	30	1,500	1,500	1,000 to 1,750
CDS	72	3,000	3,000	900
CDD	72	6,000	6,000	900
SPD	20	5,000	5,000	500 to 1,000
MSP[a]	30	3,000	10,000	750 to 1,350
STN	30	5,000	5,000	900

[a]MSP also had a delay between the stimulus and the response grids of 5,000 milliseconds.

respectively). Similarly, most cognitive tests correlated moderately with each other, with the exception of CDD, which did not correlate with the other cognitive measures, likely tapping into a different construct than the other measures. In addition, in this nonclinical population, no cognitive tests correlated with any psychological tests. This is to be expected given that the great majority of scores on the psychological tests were well within the "normal" range, and the cognitive scores were closely grouped. Finally, CES did not correlate with any measure in this nonclinical population of mostly noncombat deployed service members.

As can be seen in Table III, DANA compares favorably to existing NCATs in terms of median/mean RT (or % accuracy for CDD) and SD. The CVs are also consistent with CVs reported from ANAM data collected in 2006, 2008, and 2012 cohorts. This suggests that differences in absolute values for RT are most likely due to the testing instrument (mouse click versus stylus) rather than anything implicit in the test itself. Taken together, DANA appears to have adequate reliability and test validity in a sample of nonclinical service members across services and environments. DANA is currently being assessed in postdeployment and clinical samples.

The results reported here suggest that the DANA has promise as a next generation NCAT. Benefits of the DANA include that it (a) includes relevant psychological tests as well as standard cognitive tests; (b) is built on a portable OS with more precise timing than previous NCATs; and (c) is a portable handheld device, which is more versatile than a laptop computer. Future studies of DANA are planned to assess DANA's ability to assist frontline providers to more rapidly and accurately evaluate service members to determine the need for higher level evaluation, treatment, or readiness to return to duty.

APPENDIX

Table AI describes the eight cognitive tests and seven psychological questionnaires that were selected for the DANA test batteries. Each individual subtest parameters are shown in Table AII including stimulus presentation duration, response interval, and interstimulus interval.

ACKNOWLEDGMENTS

We thank the following people for their help in data collection and/or analysis: James Drane, Jon Farris, and Lindsay Long. This work was supported by the U.S. Navy Bureau of Medicine and Surgery, Wounded, Ill and Injured Directorate, Contract No. N00189-10-P-Z967 as well as with resources and the use of facilities at the Department of Veterans Affairs, National Center for PTSD, Pacific Islands Health Care System.

REFERENCES

1. Cernich AN, Brennan DM, Barker LM, Bleiberg J: Sources of error in computerized neuropsychological assessment. Arch Clin Neuropsychol 2007; 22S: S39–48.
2. Wilson AJ, Mellinger J, Schalk G, Williams JC: A procedure for measuring latencies in brain-computer interfaces. IEEE Trans Biomed Eng 2010; 57(7): 1785–97.
3. Weir JP: Quantifying test-retest reliability using the intraclass correlation coefficient and the SEM. J Strength Cond Res 2005; 19(1): 231–40.
4. Marx RG, Menezes A, Horovitz L, Jones EC, Warren RF: A comparison of two time intervals for test-retest reliability of health status instruments. J Clin Epidemiol 2003; 56: 730–5.
5. Reeves DL, Bleiberg J, Roebuck-Spencer T, et al: Reference values for performance on the Automated Neuropsychological Assessment Metrics V3.0 in an active duty military sample. Mil Med 2006; 171: 982–94.
6. Vincent AS, Bleiberg J, Yan S, et al: Reference data from the automated Neuropsychological Assessment Metrics for use in traumatic brain injury in an active duty military sample. Mil Med 2008; 173: 836–52.
7. Vincent AS, Roebuck-Spencer T, Gilliland K, Schlegel R: Automated Neuropsychological Assessment Metrics (v4) Traumatic Brain Injury Battery: military normative data. Mil Med 2012; 177(3): 256–69.
8. Beharelle AR, Tisserand D, Stuss DT, McIntosh AR, Levine B: Brain activity patterns uniquely supporting visual feature integration after traumatic brain injury. Front Hum Neurosci 2011; 5: 164.
9. Cernich A, Reeves D, Sun W, Bleiberg J: Automated Neuropsychological Assessment Metrics sports medicine battery. Arch Clin Neuropsychol 2007; 22S: S101–14.
10. Segalowitz SJ, Mahaney P, Santesso DL, MacGregor L, Dywan J, Willer B: Retest reliability in adolescents of a computerized neuropsychological battery used to assess recovery from concussion. NeuroRehabilitation 2007; 22(3): 243–51.
11. Bastien CH, Vallières A, Morin CM: Validation of the Insomnia Severity Index as an outcome measure for insomnia research. Sleep Med 2001; 2(4): 297–307.
12. Prins A, Ouimette P, Kimerling R, et al: The Primary Care PTSD screen (PC-PTSD). Primary Care Psychiatry 2003; 9: 9–14.
13. Prins A, Ouimette P, Kimerling R, et al: The Primary Care PTSD screen (PC-PTSD): Corrigendum. Primary Care Psychiatry 2004; 9: 151.
14. Kroenke K, Strine TW, Spitzer RL, Williams JB, Berry JT, Mokdad AH: The PHQ-8 as a measure of current depression in the general population. J Affect Disord 2009; 114(1–3): 163–73.
15. Buysse DJ, Reynolds CF III, Monk TH, Berman SR, Kupfer DJ: The Pittsburgh Sleep Quality Index: a new instrument for psychiatric practice and research. Psychiatry Res 1989; 28(2): 193–213.
16. Keane T, Fairbank J, Caddell J, Zimering R, Taylor K, Mora C: Clinical evaluation of a measure to assess combat exposure. Psychol Assessment 1989; 1: 53–5.
17. Bliese PD, Wright KM, Adler AB, Cabrera O, Castrol CA, Hoge CW: Validating the primary care posttraumatic stress disorder screen and the posttraumatic stress disorder checklist with soldiers returning from combat. J Consult Clin Psychol 2008; 76: 272–81.
18. Meterko M, Baker E, Stolzmann KL, Hendricks AM, Cicerone KD, Lew HL: Psychometric assessment of the Neurobehavioral Symptom Inventory-22: the structure of persistent postconcussive symptoms following deployment-related mild traumatic brain injury among veterans. J Head Trauma Rehabil 2012; 27(1): 55–62.