APPLIED NEUROPSYCHOLOGY: ADULT
2017, VOL. 24, NO. 6, 566–576
http://dx.doi.org/10.1080/23279095.2016.1213263

Normative data for evaluating mild traumatic brain injury with a handheld neurocognitive assessment tool

Andrea S. Vincent[a], Christopher M. Bailey[b], Charles Cowan[c], Eugenia Cox-Fuenzalida[a], Jeff Dyche[c], Kim A. Gorgens[d], Daniel C. Krawczyk[e] and Leanne Young[e]

[a]Cognitive Science Research Center, University of Oklahoma, Norman, Oklahoma, USA; [b]University Hospitals Case Medical Center, Cleveland, Ohio, USA; [c]Department of Psychology, James Madison University, Harrisonburg, Virginia, USA; [d]Graduate School of Professional Psychology, University of Denver, Denver, Colorado, USA; [e]School of Behavioral and Brain Sciences, The University of Texas at Dallas, Dallas, Texas, USA

ABSTRACT

The BrainScope® Ahead® 300 is designed for use by health care professionals to aid in the assessment of patients suspected of a mild traumatic brain injury. The purpose of the current study was to establish normative data for the cognitive test component of the Ahead 300 system and to evaluate the role of demographic factors on test performance. Healthy, community-dwelling adults between the ages of 18 and 80 recruited from five geographically distributed sites were administered Android versions of the ANAM Matching to Sample and Procedural Reaction Time tests that comprise the cognitive test component of the Ahead 300 system by trained personnel. Scores were correlated with age, education, and race. Age accounted for the majority of the variance in test scores with additional significant, but minor, contributions of education and race. Gender did not account for a significant proportion of the variance for either test. Based on these results, the normative data for 551 individuals are presented stratified by age. These are the first available normative data for these tests when administered using the Ahead 300 system and will assist health care professionals in determining the degree to which scores on the cognitive tests reflect impaired performance.

KEYWORDS

Computerized testing; concussion; handheld computers; neuropsychology; normative study

In recent years, sport and nonsport concussion/mild traumatic brain injury (mTBI) has received much attention as a public health concern, with good reason. Approximately 1.7 million concussions per year lead to emergency department (ED) visits (Centers for Disease Control and Prevention, 2003; Faul, Xu, Wald, & Coronado, 2010). In 2007, the CDC reported that 200,000 sports-related head injuries are treated in the ED annually within the United States and that sports-related concussion accounts for approximately 20% of all TBI ED visits per year (Centers for Disease Control and Prevention, 2007). Since 2007, ED visits for head trauma have dramatically increased, with particular increases in sport-related concussion (Gaw & Zonfrillo, 2016). Though the increase in visits from sport concussion was substantial, the largest increases in ED visits for head trauma between 2007 and 2011 were in children younger than 11 and adults older than 65, and were not necessarily related to sports. Concussion has been all too common in the military context as well, with blast-induced mTBI having been described as the signature injury of the Afghanistan and Iraq wars (Warden, 2006). Estimates indicate that 15–23% of service members deployed in these regions have suffered an mTBI (Hoge et al., 2008; MacGregor et al., 2010). Across contexts, mTBI represents 75% of all hospital visits associated with TBI, with estimates of the costs to the United States associated with concussion/mTBI reaching approximately $17 billion per year in medical care, lost productivity, and litigation (Centers for Disease Control and Prevention, 2003).

Neuropsychological evaluation has been described as the "cornerstone" of concussion management (Aubry et al., 2002), with multiple publications demonstrating good validity and utility in the identification of residual symptoms of concussion in both sports (Belanger & Vanderploeg, 2005; Echemendia et al., 2013; Iverson & Schatz, 2015; Nelson et al., 2016) and nonsports populations (Belanger & Vanderploeg, 2005; Karr, Areshenkoff, & Garcia-Barrera, 2014; Vanderploeg, Belanger, & Curtiss, 2009), particularly in the initial days following the injury. Computerized neurocognitive assessment tools (CNATs) have become commonplace, both within the sports arena (Covassin, Elbin,

CONTACT Andrea S. Vincent avincent@ou.edu Cognitive Science Research Center, University of Oklahoma, 350 David L Boren Blvd, Suite 1910, Norman, OK 73019, USA.

Stiller-Ostrowski, & Kontos, 2009) and outside of sports (Karr et al., 2014). In fact, CNATs have demonstrated the highest sensitivity to concussion among commonly utilized evaluation techniques (i.e., measures of postural stability and symptom endorsement) in the early days following injury (Broglio, Macciocchi, & Ferrara, 2007) as well as sensitivity and utility for the management of concussion across injury context (McCrory et al., 2013; Norris, Carr, Herzig, Labrie, & Sams, 2013).

The Automated Neuropsychological Assessment Metrics (ANAM)[1] is a computerized cognitive test battery that has been widely used to measure cognitive effects of injury, exposure, or illness (CSRC, 2013). ANAM includes a library of tests designed to be sensitive measures of attention, processing speed, working memory, and cognitive efficiency (CSRC, 2013). ANAM has demonstrated clinical utility in the evaluation and monitoring of mTBI populations in sports (Cernich, Reeves, Sun, & Bleiberg, 2007; Nelson et al., 2016), military blast injury (Bryan & Hernandez, 2012; Norris et al., 2013), and mixed clinical populations (Kane, Roebuck-Spencer, Short, Kabat, & Wilken, 2007; Woodhouse et al., 2013). Two ANAM tests have recently been transitioned to run on an Android platform for use on mobile devices (i.e., tablets and smartphones): Matching to Sample (M2S) and Procedural Reaction Time (PRO). Both subtests have shown good validity and clinical utility in a variety of mTBI populations (Bleiberg, Garmoe, Halpern, Reeves, & Nadler, 1997; Bleiberg, Kane, Reeves, Garmoe, & Halpern, 2000; Bryan & Hernandez, 2012; Cernich et al., 2004; Cernich et al., 2007; Kelly, Coldren, Parish, Dretsch, & Russell, 2012; Luethcke, Bryan, Morrow, & Isler, 2011; Yallampalli et al., 2013). For example, Bryan and Hernandez (2012) examined M2S and PRO performance in 116 service members referred to a TBI clinic in central Iraq for TBI evaluation. In comparison to those determined to not have sustained mTBI, a greater proportion of those diagnosed with mTBI ($N = 96$) demonstrated declines in response time from baseline of more than .5 standard deviations on both tests. Similarly, Kelly et al. (2012) and Coldren, Russell, Parish, Dretsch, and Kelly (2012) observed significant declines in performance on these tests 72 hours following mTBI compared to small performance improvements (i.e., practice effects) observed among controls. Coldren et al. (2012) found that these declines resolved by 5–15 days post-injury. Luethcke et al. (2011) also found slowed response times on M2S and PRO following mTBI in service members, regardless of mechanism of injury (blast vs. nonblast), and an associated decline in accuracy on M2S of almost a full standard deviation for both blast and nonblast traumatic brain injuries.

In non-military samples, Bleiberg et al. (1997, 1998) found that patients with mTBI performed worse on PRO compared to controls, even when traditional neuropsychological tests were within the normal range. Cernich et al. (2007) also reported significant differences between people with head injuries and controls when utilizing PRO. A recent prospective, head-to-head study of the efficacy of CNATs in the evaluation of sports concussion (Nelson et al., 2016), demonstrated the sensitivity to concussion of both the M2S and PRO subtests with at least equal effect size and duration of sensitivity following concussion as other commonly used CNATs. In fact, ANAM M2S was the only CNAT index that demonstrated clinically significant sensitivity to concussion at 15 days post-injury in the concussed athlete sample.

CNAT data have traditionally been interpreted clinically by collecting post-injury data, with comparison to a pre-injury baseline reference in order to identify magnitude of decline (Echemendia et al., 2013). This has also been the case with ANAM in both sports (Cernich et al., 2007) and nonsports populations (Bryan & Hernandez, 2012; Roebuck-Spencer, Vincent, Schlegel, & Gilliland, 2013). Although this clinical practice is commonplace in the evaluation and management of sports concussion, some authors have raised concerns about the use of a serial testing model in the context of concussion management, given the introduction of additional error in models that rely on repeated testing (Echemendia et al., 2013; Iverson & Schatz, 2015), most notably by issues regarding CNAT test reliability which has been shown to be less than optimal in recent years for all commonly used CNATs (Nelson et al., 2016; Resch, McCrea, & Cullum, 2013). Even when no baseline CNAT test data are available, support exists for the clinical utility of neurocognitive data in the evaluation and management of concussion based on normative (Echemendia et al., 2012) and multivariate base rate data alone (Iverson & Webbe, 2011), including the clinical interpretation of ANAM specifically using normative data in sports (Schmidt, Register-Mihalik, Mihalik, Kerr, & Guskiewicz, 2012) and non-sport mTBI populations (Ivins et al., 2015). Thus, the purpose of the current study was (1) to evaluate the role of demographic factors for the cognitive test component of the BrainScope Ahead 300[2] system made up of the

[1]ANAM is exclusively distributed by Vista LifeSciences, Inc.

[2]The Ahead 300 is under development. At the time this manuscript was submitted the Ahead 300 had not received premarket clearance from the Food and Drug Administration.

Android version of the ANAM M2S and PRO tests, (2) to establish normative data for these tests when administered on the mobile platform, (3) to report base rates of below average performance in individuals with no subjective cognitive impairment upon self-report, and (4) to examine the consistency of the current normative data to patterns observed for the same tests when administered on a PC platform.

Methods

Participants

Participants for the normative sample were recruited from the general community at five geographically distributed sites across the U.S. (Colorado, Ohio, Oklahoma, Texas, and Virginia). Data collection was conducted simultaneously at each site following an identical research protocol. Recruitment was stratified at each site by age and gender. For age, approximately equal numbers were sought from six age groups (18–30, 31–40, 41–50, 51–60, 61–70, and 71–80) for both males and females. Data collection efforts yielded data from 624 individuals across the five test sites in this age range. Each of the five sites contributed at least 86 participants with four of the five sites contributing over 100 participants each. Participants were excluded if they met any of the following criteria: history of head injury with loss of consciousness greater than 30 minutes ($N = 18$); diagnosis of severe psychiatric illness ($N = 16$); clinically significant neurological disease ($N = 20$); or use of centrally acting medication ($N = 27$); inability to understand English sufficiently to comprehend testing instructions ($N = 0$). Additionally, observations were excluded on a test-by-test basis due to low accuracy (<56%; M2S: $N = 11$; PRO: $N = 17$) or outliers that exceeded six standard deviations from the mean reaction time (M2S: $N = 1$; PRO: $N = 2$). In healthy individuals, these scores are typically indicative of a lack of understanding of the instructions or insufficient motivation.

The final sample size after these exclusions[3] was 551 for M2S and 544 for PRO. All participants were between the ages of 18 and 80 years ($M = 47.3$, $SD = 17.7$) and English speaking (Table 1). The racial composition of the final sample was 82% white. The education level varied from 8th grade or less to a doctoral degree with the majority (59%) having minimally completed a bachelor's degree. The male to female ratio was 248–303.

[3]Totals may not add to final sample size due to individuals potentially meeting more than one exclusion criteria.

Table 1. Demographic characteristics of normative sample[a].

Age, mean (*SD*)	47.3 ± 17.7
Age group, *n* (%)	
18–30	133 (24)
31–40	87 (16)
41–50	83 (15)
51–60	83 (15)
61–70	97 (18)
71–80	68 (12)
Gender, n (%)	
Male	248 (45)
Female	303 (55)
Race, *n* (%)	
White	449 (82)
Black	26 (5)
American Indian or Alaska Native	3 (<1)
Native Hawaiian or Pacific Islander	2 (<1)
Asian	12 (2)
Hispanic or Latino	31 (6)
Other or unknown	28 (5)
Education, *n* (%)	
8th grade or less	5 (<1)
Some high school, no diploma	6 (1)
High school graduate, diploma or equivalent	43 (8)
Some college, no degree	98 (18)
Trade/technical/vocational training	21 (4)
Associate's degree	49 (9)
Bachelor's degree	177 (32)
Master's degree	100 (18)
Professional degree	13 (2)
Doctorate	35 (6)
Unknown	4 (<1)
Marital Status, *n* (%)	
Single, never married	128 (23)
Married or domestic partnership	342 (63)
Widowed	20 (4)
Divorced	54 (10)
Separated	2 (<1)

*Values vary slightly by test. Values from M2S sample ($n = 551$).

Materials and procedures

BrainScope Ahead 300

The BrainScope Ahead 300 is a portable, non-invasive, point of care device in development containing multiple, selectable test modules intended to provide a configurable panel of measures supporting a multidimensional clinical evaluation for the assessment of the full spectrum of traumatic brain injury (TBI). At its core the Ahead 300 records and automatically analyzes the electroencephalogram (EEG) acquired from the frontal and frontotemporal regions, applies a classification algorithm to evaluate the likelihood of the presence of TBI visible on CT, and derives an overall measure of brain function impairment, *the EEG Brain Function Index,* expressed as a percentile relative to a large normal/uninjured population. The two cognitive performance tasks normed in this study are available in the Ahead 300, as part of the full set of options available for the multidimensional assessment of TBI.

Cognitive tests

For the current study, participants were administered a mobile version of two ANAM tests designed for use on

Table 2. Cognitive test descriptions in order of administration.

Test name	Description
Procedural Reaction Time	Measures attention and processing speed by having the user respond as quickly as possible to different sets of stimuli based on simple rules. (32 trials)
Matching to Sample	Measures visual spatial discrimination and working memory by presenting the user with a visual pattern for a specified period of time and then, following a brief delay, asking the user to select the previously seen pattern from two choices. (20 trials)

the BrainScope Ahead 300: Procedural Reaction Time (PRO) and Matching to Sample (M2S). These two tests measure reaction time, processing speed, attention, visuo-spatial processing, and working memory. Brief descriptions of each test are provided in Table 2 in the sequence of administration. The mobile version of the tests was programmed for the Android 4.1 operating system with API level 16 and was administered on the Ahead 300 investigational device.

The study was approved by the Institutional Review Board of each of the study sites in accordance with the ethical standards laid down in the Declaration of Helsinki. Each subject signed a consent form prior to their participation in the study and received financial compensation.

Test administration procedures

Following consent and study enrollment, the participants completed a short demographics questionnaire and then were administered the cognitive tests by trained test proctors. The Ahead 300 is designed for use by health care professionals to aid in the assessment of patients suspected of mTBI and includes measures of brain electrical activity (Prichep et al., 2012) and a cognitive testing component.

The tests were administered on a standard hardware configuration at all sites (Trimble Juno T41 rugged handheld computer with a 4.3" display), which will be part of the Ahead 300 system. Prior to testing, each of the participants was directed to a quiet, distraction-free environment and allowed to be seated or to stand. The participants responded to the cognitive test stimuli by tapping the screen using the thumb of each hand, as appropriate. Each test began with practice items to assist with learning procedures and minimize initial practice effects before data collection. In the event that a participant did not understand the instructions, proctors were present to provide clarification.

Device technical evaluation

While not true of all CNATs, many include response times (RTs) as either primary or secondary outcome measures. RT is defined as the latency between stimulus display and an examinee's response. While RT is often one of the most sensitive measures of cognitive impairment as a function of disease, injury, or other risk factors, it is also highly susceptible to error introduced by characteristics of the hardware or operating system environment. The rise in computerized testing coupled with the broad availability of and variability in equipment available to conduct such testing makes quantifying the RT precision and accuracy of CNATs critical. Various hardware and software technologies introduce small (ms) delays in the recording of response times. As long as these delays are consistent and repeatable, they do not negatively impact the measurement of response times for individuals or normative results because they can be factored out of the analysis.

Therefore, during the cognitive test development, response timing characteristics were evaluated for the Android versions of the two tests when delivered on the selected hardware (Trimble Juno T41). Accuracy and consistency of reaction time measurement was assessed using the Black Box Toolkit, Version 2 (Black Box Toolkit Ltd, Sheffield, UK). The Black Box Toolkit (BBTK) is a commercial off-the-shelf product used to assess the variability of device hardware and software and the resulting contributions to the measured response times. The BBTK offers the same functionality as a binary state 20 channel digital oscilloscope or logic analyzer. A robotic accuator arm was used to deliver a response on the touchscreen at a pre-specified time following stimulus presentation. In comparison to the traditional PC implementation of the same tests, additional timing latencies were observed which could be attributed to any number of factors including the touchscreen, OS, etc. More importantly, results demonstrated sufficient consistency of RT measurement with the standard deviation of the measured reaction times being between 5 and 8 milliseconds on the selected device.

Statistical analyses

All data analyses for this paper were generated using SAS software, Version 9.2 of the SAS System for Windows (SAS Institute Inc., Cary, NC). Demographic characteristics and test scores were summarized with means and standard deviations (SD) for the continuous measures and frequencies and percentages for the categorical variables. Pearson's or point-biserial correlations

(r) were used to explore the association of demographic characteristics with test performance.

Based on previous work, the primary outcome measure was throughput. Throughput is considered a measure of effectiveness or cognitive efficiency and is a combination of reaction time and accuracy (Thorne, 2006). Throughput units are reported as correct responses per minute of available response time. Higher values are indicative of better performance. Normative data are presented for Throughput, as well as mean reaction time for correct responses and percent correct.

The Throughput scores for each of the tests were initially examined for normality of distribution. Multiple regression analysis was conducted to examine the effects of age, gender (0 = male; 1 = female), education (0 = high school diploma or less; 1 = some college or more), and race (0 = white; 1 = other) on test scores. For the regression analyses, age was treated as a continuous variable (i.e., not categorized into discrete age groups). The education and racial groupings were made on the basis of the availability of data as well as a preliminary analysis looking for natural break points among the response categories.

Prior to compilation of the normative data, analysis of variance using the general linear model was conducted to examine differences according to age (6 groups), education (2 groups), and race (2 groups). Effect sizes were calculated using the generalized ω^2 (omega-squared) statistic (small = 0.01, medium = 0.06, large = 0.15) (Olejnik & Algina, 2003). Percentile ranks, means, and standard deviations were computed for the overall normative sample and for relevant subgroups. Percentiles were calculated using the UNIVARIATE procedure in SAS using the empirical distribution function with averaging method. Base rates of "Below Average" (throughput between 2nd and 9th percentile) and "Clearly Below Average" (throughput at or below 2nd percentile) performance were calculated for each test. Base rate calculations were based on Throughput in comparison to age- and gender-matched data.

Linear regression analysis was used to examine the consistency of the slope of the regression lines describing the association between performance and relevant demographics for the current data in comparison to existing normative data for the tests when administered on a standard PC platform.

Results

Multiple regression analysis was used to examine the contributions of age, gender, education, and race to Throughput scores from each of the tests. None of the interaction effects (2- or 3-way) were significant, thus, they were not included in the final models. Regression coefficients and zero-order correlations (Pearson for age and point biserial for gender, education, and race) are presented in Table 3. For both M2S and PRO, significant zero-order correlations were observed between Throughput and age, education, and race. While significant, the correlations with education and race were small in magnitude. These three variables also had significant ($p < .05$) partial effects in the regression model. Age accounted for the largest proportion of the variance with lesser, but significant, effects observed for both education and race. Gender did not account for a significant proportion of the variance for either test. The four predictor model was able to account for 28% of the variance in M2S scores, $F(4, 545) = 52.62$, $p < .0001$, $R^2 = .28$ and 35% of the variance in PRO scores, $F(4, 538) = 70.91$, $p < .0001$, $R^2 = .35$.

Prior to compilation of the stratified normative data, a three-way ANOVA was conducted for each test to assess effects of the potential stratification variables identified from the multiple regression. Age was grouped

Table 3. Throughput scores related to demographic characteristics.

	Zero-Order r							
Variable	Race	Education	Gender	Age	TP	β	b	$SE\ b$
Matching to Sample ($N = 551$)								
Age					−.47**	−.51**	−.35	.03
Gender (0 = M; 1 = F)				−.00	−.04	−.04	−1.07	.90
Education (0 = HS or less; 1 = Some college or more)			.01	.03	.14**	.13**	5.19	1.53
Race (0 = White; 1 = Other)		−.19**	.02	−.20**	−.11**	−.19**	−5.84	1.19
								Intercept = 47.50
	$R^2 = .28$							
Procedural Reaction Time ($N = 543$)								
Age					−.51**	−.55**	−0.75	0.05
Gender (0 = M; 1 = F)				.01	−.07	−.06	−3.09	1.67
Education (0 = HS or less; 1 = Some college or more)			−.02	.04	.21**	.19**	14.32	2.73
Race (0 = White; 1 = Other)		−.21**	.01	−.20**	−.12**	−.19	−11.32	2.22
								Intercept = 126.24
	$R^2 = .35$							

Note. TP = Throughput.
$^*p < .05$. $^{**}p < .01$.

Table 4. Analysis of variance source table and effect sizes for age, education, and race effects on throughput scores.

Test	Effect	*df*	*F*	*p*	ω^2
Matching to Sample	Age	2	78.60	<.0001	.22
	Education	1	10.66	.0010	.02
	Race	1	18.68	<.0001	.03
Procedural Reaction Time	Age	3	72.00	<.0001	.28
	Education	1	24.46	<.0001	.04
	Race	1	23.38	<.0001	.04

Note. Generalized ω^2 effect size estimates reported to 2 digits. Rule of Thumb: 0.01 is small, 0.06 is medium, 0.15 is large.

in a manner consistent with the groupings used for recruitment: 18–30, 31–40, 41–50, 51–60, 61–70, and 71–80 years. The education and race groups were as previously defined. Significant main effects were observed for all three variables for both tests (M2S: $F(4, 546) = 44.21$, $p < .0001$, $R^2 = .24$; PRO: $F(5, 538) = 50.72$, $p < .0001$, $R^2 = .32$). Mean cognitive efficiency was observed to decline with age and was higher among those with higher education and whites. The largest proportion of the variance in Throughput scores was explained by age (22% for M2S and 28% for PRO) with education and race each explaining 3% or less for M2S and 4% or less each for PRO (Table 4).

Due to the relatively small additional contribution of the education and race variables and the resulting small sample sizes when all three stratification variables are included, the normative data are presented in Tables 5 and 6 stratified only by age. Normative tables include the mean, SD, minimum (0%), and maximum (100%) percentile scores, first (25%), second (50%), and third

Table 5. Mean, standard deviation, and percentile scores for matching to sample test.

				Percentile								
Variable	Age	n	Mean ± SD	0	2	9	25	50	75	91	98	100
Mean RT	All	551	1744 ± 651	5303	3460	2706	2076	1586	1276	1075	821	724
	18–30	133	1415 ± 381	2552	2419	2022	1634	1316	1178	934	799	724
	31–40	87	1535 ± 510	2979	2861	2492	1772	1420	1187	947	763	756
	41–50	83	1685 ± 582	3467	3143	2565	2046	1539	1243	1080	817	755
	51–60	83	1856 ± 555	3460	3448	2663	2199	1809	1480	1096	884	765
	61–70	97	2029 ± 711	4302	4148	2977	2352	1901	1489	1249	1045	1020
	71–80	68	2188 ± 856	5303	5189	3280	2679	1974	1600	1384	1172	1056
Correct (%)	All	551	89.6 ± 8.2	60	70	75	85	90	95	100	100	100
	18–30	133	92.2 ± 6.6	70	75	85	90	95	95	100	100	100
	31–40	87	90 ± 7.4	65	70	80	85	90	95	100	100	100
	41–50	83	91.2 ± 9	60	65	80	85	90	100	100	100	100
	51–60	83	88.4 ± 7.8	70	70	75	85	90	95	100	100	100
	61–70	97	88 ± 8.4	65	70	75	85	90	95	100	100	100
	71–80	68	86.4 ± 9	60	65	75	80	87.6	95	95	100	100
Throughput	All	551	33.8 ± 12.2	10.2	14.2	18.8	24.2	32.2	41.4	49.8	63.4	75.8
	18–30	133	41 ± 11.4	20	22.6	26.4	32.2	41	47.2	58.2	70.2	73
	31–40	87	37.8 ± 12.2	12.8	18.6	20.8	29.2	37.4	44.8	52.6	70	75.8
	41–50	83	35.2 ± 12.6	12	14.2	19.6	25.8	35.4	42.6	53.8	61.8	69.6
	51–60	83	30.2 ± 10.8	12.8	13.6	19.4	22	27.4	35.8	47	57.4	62.6
	61–70	97	28 ± 9.2	12.2	13.4	16	20.2	26.4	34	42.2	48.8	49.4
	71–80	68	25.4 ± 8	10.2	11.4	15	18.8	25.4	31	35.6	40.2	43.8

Table 6. Mean, standard deviation, and percentile scores for procedural reaction time test.

				Percentile								
Variable	Age	*n*	Mean ± *SD*	0	2	9	25	50	75	91	98	100
Mean RT	All	544	625 ± 204	2036	1354	848	683	579	499	448	412	364
	18–30	128	517 ± 83	777	767	622	557	504	463	414	392	364
	31–40	90	564 ± 141	1453	909	724	597	526	477	448	426	422
	41–50	83	604 ± 212	1658	1477	848	644	544	479	441	424	413
	51–60	81	682 ± 203	1494	1369	918	759	633	550	498	412	411
	61–70	95	677 ± 163	1440	1249	901	727	643	578	531	446	431
	71–80	67	797 ± 306	2036	1893	1167	847	722	623	546	497	480
Correct (%)	All	544	96.6 ± 5.4	56.2	78.2	90.6	96.8	96.8	100.0	100.0	100.0	100.0
	18–30	128	96.6 ± 4.6	71.8	78.2	90.6	93.8	96.8	100.0	100.0	100.0	100.0
	31–40	90	97.4 ± 3.2	84.4	90.6	93.8	96.8	96.8	100.0	100.0	100.0	100.0
	41–50	83	96.6 ± 5.8	56.2	87.6	90.6	93.8	96.8	100.0	100.0	100.0	100.0
	51–60	81	96.4 ± 7.6	56.2	68.8	90.6	96.8	100.0	100.0	100.0	100.0	100.0
	61–70	95	96.4 ± 6	71.8	71.8	87.6	96.8	96.8	100.0	100.0	100.0	100.0
	71–80	67	97 ± 5	65.6	84.4	93.8	96.8	96.8	100.0	100.0	100.0	100.0
Throughput	All	544	99.8 ± 23.8	24.0	40.2	67.8	85.0	101.4	116.6	129.2	140.8	165.0
	18–30	128	115 ± 17.8	73.8	75.8	88.0	105.0	115.6	127.0	138.2	146.6	165.0
	31–40	90	107.8 ± 19.4	41.2	58.0	77.4	97.0	111.0	122.6	129.8	136.4	137.8
	41–50	83	103.4 ± 25	36.2	40.6	63.2	90.2	107.6	122.0	133.6	139.4	141.2
	51–60	81	91.2 ± 23.8	25.4	27.0	65.4	77.6	94.2	106.4	120.4	140.8	145.8
	61–70	95	89.2 ± 18.6	24.0	33.6	63.6	79.8	91.8	100.4	110.8	117.4	120.8
	71–80	67	80.4 ± 20.6	26.2	29.4	48.6	69.4	81.6	94.2	106.8	117.0	119.8

Table 7. Proportion of healthy individuals obtaining "Below Average" or "Clearly Below Average" scores on the Ahead 300 cognitive tests.

Test	N	"Below Average"	"Clearly Below Average"
Matching to Sample	551	7.6	2.7
Procedural Reaction Time	544	7.4	2.8

Note. "Below average" includes scores falling at or below the 9th percentile and above the 2nd percentile. "Clearly below average" includes scores at or below the 2nd percentile.

(75%) quartile scores, and 2nd, 9th, 91st, and 98th percentile scores for each subgroup.

Approximately 10% of individuals scored in the Below Average or Clearly Below Average range on each test despite no objective cognitive impairment (Table 7). A total of 16.6% of individuals earned at least one score in the Below Average or Clearly Below Average range. The majority of these (81%) scored in the impaired range on only one of the tests. Of this presumed healthy sample, only 17 (3%) had impaired scores (either Below or Clearly Below Average) on both tests (Table 8).

The pattern of age-related change observed for data collected on the Ahead 300 was compared to existing normative data ($N = 419$) for the same tests when administered on a PC for a comparable age range (ages 18–80) (CSRC, 2013). There was no difference in the slope of the regression line of age in predicting Throughput scores for either of the tests (M2S: $b = -0.337$, SEM $= 0.03$ vs. $b = -0.325$, SEM $= 0.03$, for PC and mobile platforms, respectively, $p = 0.65$; PRO: $b = -0.605$, SEM $= 0.05$ vs. $b = -0.689$, SEM $= 0.05$, for PC and mobile platforms, respectively, $p = 0.09$). Table 9 presents the mean difference in Throughput scores for sequential age groups that were observed in data obtained on both platforms. Consistent with the regression analysis, declines in Throughput scores were observed with increasing age group for all data obtained on the PC platform. This age-related decline was also observed for data on the mobile platform. The magnitudes of the declines were also comparable.

Table 8. Frequency of impaired scores on the Ahead 300 cognitive tests.

Test	N (%)
2 scores Clearly Below Average	3 (0.6)
1 score Clearly Below Average AND 1 score Below Average	9 (1.7)
1 score Clearly Below Average AND 1 score Average	13 (2.4)
2 scores Below Average	5 (0.9)
1 score Below Average AND 1 score Average	59 (11.0)
2 scores Average	447 (83.3)

Note. Includes only those individuals with scores on both tests ($N = 536$). "Average" (or above) includes scores above the 9th percentile. "Below average" includes scores falling at or below the 9th percentile and above the 2nd percentile. "Clearly below average" includes scores at or below the 2nd percentile.

Table 9. Mean throughput difference (95% CI) by age group for mobile and PC platforms.

Age group (yrs)	PC Δ_{2-1}	Mobile Δ_{2-1}
	Matching to sample	
18–30 (1): 31–40 (2)	−4.5 (−0.2, 9.2)	−3.1 (−1.2, 7.4)
31–40 (1): 41–50 (2)	−2.4 (−3.2, 8.0)	−2.5 (−2.3, 7.3)
41–50 (1): 51–60 (2)	−4.3 (−1.2, 9.8)	−5.1 (0.3, 10.0)
51–60 (1): 61–70 (2)	−1.8 (−5.4, 9.0)	−2.2 (−2.5, 6.9)
61–70 (1): 71–80 (2)	−1.7 (−7.9, 11.3)	−2.5 (−2.4, 7.5)
	Procedural reaction time	
18–30 (1): 31–40 (2)	−7.4 (−0.1, 14.8)	−7.1 (−1.1, 15.2)
31–40 (1): 41–50 (2)	−1.0 (−7.9, 9.8)	−4.4 (−4.6, 13.4)
41–50 (1): 51–60 (2)	−9.6 (1.1, 18.2)	−12.2 (2.9, 21.4)
51–60 (1): 61–70 (2)	−6.4 (−4.8, 17.6)	−2.1 (−6.9, 11.0)
61–70 (1): 71–80 (2)	−12.9 (−1.7, 27.5)	−8.8 (−0.7, 18.2)

Discussion

The goal of this study was to establish normative data for the cognitive test component of the BrainScope Ahead 300 system and to examine the effects of various demographic factors, including age, gender, race, and education, on test performance. Normative data are especially important in this case in order to validate the mobile methodology for data collection and to establish the feasibility of use in general community populations.

Valid interpretation of neuropsychological test scores necessitates information about typical performance so that deviations from this can be identified. The absence of such normative data can lead to faulty conclusions about the clinical meaning of test results. Traditional instruments are sensitive to user error, and face-to-face interviews with allied health professionals are not diagnostic, typically having only a limited role in medical settings (Malhotra et al., 2015; Persoon, Van der Cruijsen, Schlattmann, Simmes, & Van Achterberg, 2011). In the absence of neuropsychological services, computerized screening tests are well suited for use in brief clinical evaluations of the common causes of neurocognitive impairment including attention deficit/hyperactivity disorder, traumatic brain injury, and dementia (Gualtieri & Johnson, 2006).

The associations of test performance with age, education, and race are broadly consistent with previous research examining cognitive performance in healthy populations. The major finding indicates a general decline in performance with age in the form of decreases in the Throughput measure. This effect was minimal among the younger age groups which is consistent with prior research showing that age-related cognitive changes are typically small for individuals 20–40 years old (Craik & Bialystok, 2006). Declines in Throughput can result from slowing reaction times, decreasing accuracy, or both. In this case, a closer examination of the component reaction time and accuracy scores reveals that the observed differences were the result of faster reaction times among younger individuals, those with higher education, and among whites rather than any differences in accuracy of responding. This is not unexpected as many healthy individuals will perform at near ceiling-effects for accuracy on PRO and M2S when cognition is not compromised. This pattern of results is consistent with prior findings from standard neuropsychological tests showing that psychomotor speed tends to slow with increasing age (Heaton, Ryan, Grant, & Matthews, 1996). Results are also consistent with previously reported ANAM data (Roebuck-Spencer et al., 2008; Vincent et al., 2008; Vincent, Roebuck-Spencer, Gilliland, & Schlegel, 2012).

While age was a factor affecting M2S and PRO performance, we found no significant effects of gender for either of the tests in the current study. This finding differs from prior research on the ANAM M2S and PRO tests in military samples (Vincent et al., 2008; Vincent et al., 2012). Those prior studies had much larger sample sizes and may have been overpowered to detect differences associated with gender.

We anticipated differences in the timing accuracy achieved on the standard PC platforms in comparison to the Android mobile device, and differences due to response modality of the current implementation (touchscreen using thumbs on the mobile device in comparison to mouse button responses on the PC). Specifically, we expected to find absolute differences in the reaction times measured in the current study in comparison to previously collected data on a standard PC platform. However, despite these differences, we expected that the *pattern* of performance on the current handheld device would be comparable to data obtained from these tests on a standard PC platform due to the similarities in the test characteristics and presentation (which were identical). These data provide preliminary evidence to demonstrate that despite being a new platform, the pattern of results obtained on this new platform is likely to be consistent with previously studies showing differences in performance related to a number of risk factors, including concussion. Future research should explore this further in clinical samples. Additionally, reliability and validity studies are still needed to fully document the psychometric properties of these tests on the mobile platform.

The observed base rates suggest that a subset of presumed healthy individuals may perform at an impaired level on either the PRO and M2S tests. Despite using slightly different cut-points, these rates are comparable to those reported by Vincent et al. (2012) for the standard ANAM version of these tests when administered in a military sample. In that study, 7% of service members scored more than 1.3 SD below the mean for M2S and 9% for PRO. The current data further suggest that it is rare for a healthy individual to score in an impaired range on *both* tests, especially when the most stringent criteria of scores at or below the 2nd percentile is applied. Overall, these data are consistent with base rates reported for other test batteries (Heaton et al., 1996; Schretlen, Testa, Winicki, Pearlson, & Gordon, 2008), although rates will vary depending on the number of tests in a battery. By documenting the prevalence of impaired scores among healthy individuals, clinicians can make more informed decisions regarding the scores observed in individual patients which may help to avoid over-diagnosis of clinical impairment.

Standardized administration is an important component of cognitive testing. Cognitive testing on traditional PC platforms typically involves administration in an indoor setting, in a seated position, with a computer screen in front of the examinee. The use of mobile devices presents challenges to these standardized procedures due to the ease of use in a variety of environments. Test environment, particularly one with significant distractions, can influence performance. For the current study, participants were tested in the community in a variety of settings. However, in each setting the participant was directed to a quiet, distraction-free environment. The participant was allowed to be seated or to stand. A proctor monitored the participant during the entire testing period to ensure proper use of the device and to be available to answer questions.

Norms should reflect a broad range of ages and educational levels with representation of the diversity of the populations intended for assessment. The representativeness of this sample is limited by race where 82% ($N = 449$) of the sample describe themselves as White. The underrepresentation of racial diversity does not allow the sample to be stratified by ethnicity within each age or education bracket. Additionally, the sites were selected to provide regional diversity, but not all geographic regions of the United States were sampled. Importantly, because the data were collected by research personnel from universities and research centers, the normative sample was largely drawn from urban areas. The proximity to academic centers also likely impacted the educational diversity in the sample, here only 11 people reported earning less than a high school diploma (<2%). Studies have shown that literacy influences the brain's organization of cognition (Matute et al., 2012). The impact is likely to be larger on more difficult test batteries, but the impact of literacy and education on cognitive function cannot be anticipated from these data. Research also suggests a complex relationship between education and the cognitive ability associated with age. The current sample shows good representation of ages from 18 to 80 but the interaction between age-associated changes and education may be heterogeneous and warrants additional research attention (de Azeredo Passos et al., 2015). All told, the expansion of the database to include more variability in education and more minority Americans will make it possible to address that issue and to standardize scores within narrower age parameters.

In summary, tools that allow the timely evaluation of a broad range of cognitive functions are a priority for research and practice. The most promising instruments should undergo this type of rigorous standardization and psychometric testing so that they can demonstrate their clinical utility prior to their deployment into settings for use by medical personnel. The normative data and associated base rates presented here will aid medical professionals in the evaluation and treatment in medical populations suspected of having sustained mTBI. The cognitive test component of the BrainScope Ahead 300 is the only available hand-held CNAT to provide norms across a broad range of ages and educational levels that can address this need.

Funding

Funding for this study was provided by BrainScope Company, Inc., Bethesda, MD.

References

Aubry, M., Cantu, R., Dvorak, J., Graf-Baumann, T., Johnston, K., Kelly, J., … Schamasch, P. (2002). Summary and agreement statement of the First International Conference on Concussion in Sport, Vienna 2001. Recommendations for the improvement of safety and health of athletes who may suffer concussive injuries. *British Journal of Sports Medicine*, *36*(1), 6–10. doi:10.1136/bjsm.36.1.6

Belanger, H. G., & Vanderploeg, R. D. (2005). The neuropsychological impact of sports-related concussion: A meta-analysis. *Journal of the International Neuropsychological Society*, *11*(4), 345–357. doi:10.1017/S1355617705050411

Bleiberg, J., Garmoe, W. S., Halpern, E. L., Reeves, D. L., & Nadler, J. D. (1997). Consistency of within-day and across-day performance after mild brain injury. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *10*(4), 247–253.

Bleiberg, J., Halpern, E. L., Reeves, D., & Daniel, J. C. (1998). Future directions for the neuropsychological assessment of sports concussion. *The Journal of Head Trauma Rehabilitationt*, *13*(2), 36–44.

Bleiberg, J., Kane, R. L., Reeves, D. L., Garmoe, W. S., & Halpern, E. (2000). Factor analysis of computerized and traditional tests used in mild brain injury research. *The Clinical Neuropsychologist*, *14*(3), 287–294. doi:10.1076/1385-4046(200008)14:3;1-P;FT287

Broglio, S. P., Macciocchi, S. N., & Ferrara, M. S. (2007). Sensitivity of the concussion assessment battery. *Neurosurgery*, *60*(6), 1050–1057; discussion 1057–1058. doi:10.1227/01.NEU.0000255479.90999.C0

Bryan, C., & Hernandez, A. M. (2012). Magnitudes of decline on Automated Neuropsychological Assessment Metrics subtest scores relative to predeployment baseline performance among service members evaluated for traumatic brain injury in Iraq. *Journal of Head Trauma Rehabilitation*, *27*(1), 45–54. doi:10.1097/HTR.0b013e318238f146

Centers for Disease Control and Prevention. (2003). *Report to congress on mild traumatic brain injury in the United States: Steps to prevent a serious public health problem*. Atlanta, GA: National Center for Injury Prevention and Control.

Centers for Disease Control and Prevention. (2007). Nonfatal traumatic brain injuries from sports and recreation activities - United States, 2001–2005. *Morbidity and Mortality Weekly Report*, *56*(29), 733–737. doi:10.1001/jama.298.11.1271

Cernich, A. N., Bleiberg, J., Roebuck-Spencer, T. M., Ivins, B., Schwab, K., Reeves, D. L., … Warden, D. (2004). *Prediction of concussion status using a computerized neuropsychological measure: A logistic regression analysis*. Abstract presented at the 32nd annual meeting of the International Neuropsychological Society conference, February 4–7, Baltimore, Maryland, USA.

Cernich, A. N., Reeves, D., Sun, W., & Bleiberg, J. (2007). Automated neuropsychological assessment metrics sports medicine battery. *Archives of Clinical Neuropsychology*, *22*(Suppl 1), S101–S114. doi:10.1016/j.acn.2006.10.008

Coldren, R. L., Russell, M. L., Parish, R. V., Dretsch, M., & Kelly, M. P. (2012). The ANAM lacks utility as a diagnostic or screening tool for concussion more than 10 days following injury. *Military Medicine*, *177*(2), 179–183. doi:10.7205/milmed-d-11-00278

Covassin, T., Elbin, R. J., III, Stiller-Ostrowski, J. L., & Kontos, A. P. (2009). Immediate post-concussion assessment and cognitive testing (ImPACT) practices of sports medicine professionals. *Journal of Athletic Training*, *44*(6), 639–644. doi:10.4085/1062-6050-44.6.639

Craik, F. I. M., & Bialystok, E. (2006). Cognition through the lifespan: Mechanisms of change. *Trends in Cognitive Sciences*, *10*(3), 131–138. doi:10.1016/j.tics.2006.01.007

CSRC. (2013). *ANAM technical manual*. Norman, OK: Cognitive Science Research Center, University of Oklahoma.

de Azeredo Passos, V. M., Giatti, L., Bensenor, I., Tiemeier, H., Ikram, M. A., de Figueiredo, R. C., … Barreto, S. M. (2015). Education plays a greater role than age in cognitive test performance among participants of the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). *BMC Neurology*, *15*(1), 1–9. doi:10.1186/s12883-015-0454-6

Echemendia, R. J., Bruce, J. M., Bailey, C. M., Sanders, J. F., Arnett, P., & Vargas, G. (2012). The utility of post-concussion neuropsychological data in identifying cognitive change following sports-related MTBI in the absence of baseline data. *The Clinical Neuropsychologist*, *26*(7), 1077–1091. doi:10.1080/13854046.2012.721006

Echemendia, R. J., Iverson, G. L., McCrea, M., Macciocchi, S. N., Gioia, G. A., Putukian, M., & Comper, P. (2013).

Advances in neuropsychological assessment of sport-related concussion. *British Journal of Sports Medicine*, *47* (5), 294–298. doi:10.1136/bjsports-2013-092186

Faul, M., Xu, L., Wald, M. M., & Coronado, V. G. (2010). *Traumatic brain injury in the United States: Emergency department visits, hospitalizations and deaths 2002–2006*. Atlanta, GA: Centers for Disease Control and Prevention, National Center for Injury Prevention and Control.

Gaw, C. E., & Zonfrillo, M. R. (2016). Emergency department visits for head trauma in the United States. *BMC Emergency Medicine*, *16*(1), 5. doi:10.1186/s12873-016-0071-8

Gualtieri, C. T., & Johnson, L. G. (2006). Reliability and validity of a computerized neurocognitive test battery, CNS vital signs. *Archives of Clinical Neuropsychology*, *21* (7), 623–643. doi:10.1016/j.acn.2006.05.007

Heaton, R. K., Ryan, L., Grant, I., & Matthews, C. G. (1996). Demographic influences on neuropsychological test performance. In I. Grant & K. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (2nd ed., pp. 141–163). New York, NY: Oxford University Press.

Hoge, C. W., McGurk, D., Thomas, J. L., Cox, A. L., Engel, C. C., & Castro, C. A. (2008). Mild traumatic brain injury in U.S. Soldiers returning from Iraq. *New England Journal of Medicine*, *358*(5), 453–463. doi:10.1056/NEJMoa072972

Iverson, G. L., & Schatz, P. (2015). Advanced topics in neuropsychological assessment following sport-related concussion. *Brain Injury*, *29*(2), 263–275. doi:10.3109/02699052.2014.965214

Iverson, G. L., & Webbe, F. M. (2011). Evidence-based neuropsychological assessment in sport-related concussion. In F. M. Webbe (Ed.), *The Handbook of Sport Neuropsychology* (pp. 131–153). New York, NY: Springer.

Ivins, B. J., Lange, R. T., Cole, W. R., Kane, R., Schwab, K. A., & Iverson, G. L. (2015). Using base rates of low scores to interpret the ANAM4 TBI-MIL battery following mild traumatic brain injury. *Archives of Clinical Neuropsychology*, *30*(1), 26–38. doi:10.1093/arclin/acu072

Kane, R. L., Roebuck-Spencer, T. M., Short, P., Kabat, M., & Wilken, J. (2007). Identifying and monitoring cognitive deficits in clinical populations using Automated Neuropsychological Assessment Metrics (ANAM) tests. *Archives of Clinical Neuropsychology*, *22*(S1), S115–S126. doi:10.1016/j.acn.2006.10.006

Karr, J. E., Areshenkoff, C. N., & Garcia-Barrera, M. A. (2014). The neuropsychological outcomes of concussion: A systematic review of meta-analyses on the cognitive sequelae of mild traumatic brain injury. *Neuropsychology*, *28*(3), 321–336. doi:10.1037/neu0000037

Kelly, M. P., Coldren, R. L., Parish, R. V., Dretsch, M. N., & Russell, M. L. (2012). Assessment of acute concussion in the combat environment. *Archives of Clinical Neuropsychology*, *27*(4), 375–388. doi:10.1093/arclin/acs036

Luethcke, C. A., Bryan, C. J., Morrow, C. E., & Isler, W. C. (2011). Comparison of concussive symptoms, cognitive performance, and psychological symptoms between acute blast-versus nonblast-induced mild traumatic brain injury. *Journal of the International Neuropsychological Society*, *17*(1), 36–45. doi:10.1017/S1355617710001207

MacGregor, A. J., Shaffer, R. A., Dougherty, A. L., Galarneau, M. R., Raman, R., Baker, D. G., … Corson, K. S. (2010). Prevalence and psychological correlates of traumatic brain injury in operation Iraqi freedom. *Journal of Head Trauma Rehabilitation*, *25*(1), 1–8. doi:10.1097/HTR.0b013e3181c2993d

Malhotra, R., Haaland, B. A., Chei, C.-L., Chan, A., Malhotra, C., & Matchar, D. B. (2015). Presence of and correction for interviewer error on an instrument assessing cognitive function of older adults. *Geriatrics & Gerontology International*, *15*(3), 372–380. doi:10.1111/ggi.12331

Matute, E., Montiel, T., Pinto, N., Rosselli, M., Ardila, A., & Zarabozo, D. (2012). Comparing cognitive performance in illiterate and literate children. *International Review of Education*, *58*(1), 109–127.

McCrory, P., Meeuwisse, W. H., Aubry, M., Cantu, R. C., Dvorak, J., Echemendia, R. J., … Turner, M. (2013). Consensus statement on concussion in sport: The 4th International Conference on Concussion in Sport, Zurich, November 2012. *Journal of Athletic Training*, *48*(4), 554–575. doi:10.4085/1062-6050-48.4.05

Nelson, L. D., LaRoche, A. A., Pfaller, A. Y., Lerner, E. B., Hammeke, T. A., Randolph, C., … McCrea, M. A. (2016). Prospective, head-to-head study of three computerized neurocognitive assessment tools (CNTs): Reliability and validity for the assessment of sport-related concussion. *Journal of the International Neuropsychological Society*, *22* (1), 24–37. doi:10.1017/S1355617715001101

Norris, J. N., Carr, W., Herzig, T., Labrie, D. W., & Sams, R. (2013). ANAM4 TBI reaction time-based tests have prognostic utility for acute concussion. *Military Medicine*, *178* (7), 767–774. doi:10.7205/MILMED-D-12-00493

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods*, *8*(4), 434–447. doi:10.1037/1082-989X.8.4.434

Persoon, A., Van der Cruijsen, M., Schlattmann, N., Simmes, F., & Van Achterberg, T. (2011). How older people nurses assess cognitive function through daily observation. *International Journal of Older People Nursing*, *6*(2), 102–109. doi:10.1111/j.1748-3743.2010.00234.x

Prichep, L. S., Jacquin, A., Filipenko, J., Dastidar, S. G., Zabele, S., Vodencarevic, A., & Rothman, N. S. (2012). Classification of traumatic brain injury severity using informed data reduction in a series of binary classifier algorithms. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *20*(6), 806–822. doi:10.1109/TNSRE.2012.2206609

Resch, J. E., McCrea, M. A., & Cullum, C. M. (2013). Computerized neurocognitive testing in the management of sport-related concussion: An update. *Neuropsychology Review*, *23* (4), 335–349. doi:10.1007/s11065-013-9242-5

Roebuck-Spencer, T. M., Reeves, D. L., Bleiberg, J., Cernich, A. N., Schwab, K., Ivins, B., … Warden, D. (2008). Influence of demographics on computerized cognitive testing in a military sample. *Military Psychology*, *20*, 187–203. doi:10.1080/08995600802118825

Roebuck-Spencer, T. M., Vincent, A. S., Schlegel, R. E., & Gilliland, K. (2013). Evidence for added value of baseline testing in computer-based cognitive assessment. *Journal of Athletic Training*, *48*(4), 499–505. doi:10.4085/1062-6050-48.3.11

Schmidt, J. D., Register-Mihalik, J. K., Mihalik, J. P., Kerr, Z. Y., & Guskiewicz, K. M. (2012). Identifying Impairments after concussion: Normative data versus individualized baselines. *Medicine & Science in Sports & Exercise*, *44*(9), 1621–1628. doi:10.1249/MSS.0b013e318258a9fb

Schretlen, D. J., Testa, S. M., Winicki, J. M., Pearlson, G. D., & Gordon, B. (2008). Frequency and bases of abnormal performance by healthy adults on neuropsychological testing. *Journal of the International Neuropsychological Society*, *14* (3), 436–445. doi:10.1017/S1355617708080387

Thorne, D. R. (2006). Throughput: A simple performance index with desirable characteristics. *Behavior Research Methods*, *38*(4), 569–573. doi:10.3758/bf03193886

Vanderploeg, R. D., Belanger, H. G., & Curtiss, G. (2009). Mild traumatic brain injury and posttraumatic stress disorder and their associations with health symptoms. *Archives of Physical Medicine and Rehabilitation*, *90*(7), 1084–1093. doi:10.1016/j.apmr.2009.01.023

Vincent, A. S., Bleiberg, J., Yan, S., Ivins, B., Reeves, D. L., Schwab, K., … Warden, D. (2008). Reference data from the automated Neuropsychological Assessment Metrics for use in traumatic brain injury in an active duty military sample. *Military Medicine*, *173*(9), 836–852.

Vincent, A. S., Roebuck-Spencer, T. M., Gilliland, K., & Schlegel, R. (2012). Automated neuropsychological assessment metrics (v4) traumatic brain injury battery: Military normative data. *Military Medicine*, *177*(3), 256–269. doi:10.7205/milmed-d-11-00289

Warden, D. (2006). Military TBI during the Iraq and Afghanistan wars. *Journal of Head Trauma Rehabilitation*, *21*(5), 398–402. doi:10.1097/00001199-200609000-00004

Woodhouse, J., Heyanka, D. J., Scott, J., Vincent, A., Roebuck-Spencer, T. M., Domboski-Davidson, K., … Adams, R. (2013). Efficacy of the ANAM general neuropsychological screening battery (ANAM GNS) for detecting neurocognitive impairment in a mixed clinical sample. *The Clinical Neuropsychologist*, *27*(3), 376–385. doi:10.1080/13854046.2012.762427

Yallampalli, R., Wilde, E. A., Bigler, E. D., McCauley, S. R., Hanten, G., Troyanskaya, M., … Levin, H. S. (2013). Acute white matter differences in the fornix following mild traumatic brain injury using diffusion tensor imaging. *Journal of Neuroimaging*, *23*(2), 224–227. doi:10.1111/j.1552-6569.2010.00537.x