Journal of the International Neuropsychological Society (2023), **29**, 415–420

doi:10.1017/S1355617717000856

RETRACTED - Intra-individual Cognitive Variability: An Examination of ANAM4 TBI-MIL Simple Reaction Time Data from Service Members with and without Mild Traumatic Brain Injury

Wesley R. Cole,[1,2,3] Emma Gregory,[1,3] Jacques P. Arrieux,[1,2,3] AND F. Jay Haran[4]

[1]Defense and Veterans Brain Injury Center, Silver Spring, Maryland and Fort Bragg, North Carolina
[2]Womack Army Medical Center, Fort Bragg, North Carolina
[3]General Dynamics Health Solutions, Fairfax, Virginia
[4]Uniformed Service University of the Health Sciences, Bethesda, Maryland

(RECEIVED February 24, 2017; FINAL REVISION May 26, 2017; ACCEPTED July 18, 2017; First Published online 11 September 2017)

Abstract

Objectives: The Automated Neuropsychological Assessment Metrics 4 TBI-MIL (ANAM4) is a computerized cognitive test often used in post-concussion assessments with U.S. service members (SMs). Existing evidence, however, remains mixed regarding ANAM4's ability to identify cognitive issues following mild traumatic brain injury (mTBI). Studies typically examine ANAM4 by comparing mean scores to baseline or normative scores. A more fine-grained approach involves examining inconsistency within an individual's performance. **Methods:** Data from a sample of 231 were healthy control SMs and 100 SMs within 7 days of mTBI who took the ANAM4 were included in analyses. We examine each individual's performance on a simple reaction time (SRT) subtest that is administered at the beginning (SRT1) and end (SRT2) of the ANAM4 battery, and calculate the standard deviation of difference scores by trial across administrations. **Results:** Multivariate analysis of variance and univariate analyses revealed group differences across all comparisons ($p < .001$) with pairwise comparisons revealing higher intra-individual variability and slower raw reaction time for the mTBI group compared with controls. Effect sizes were small though exceeded the recommended minimum practical effect size (ES > 0.41). **Conclusions:** While inconsistencies in performance are often viewed as noise or test error, the results suggest intra-individual cognitive variability may be more sensitive than central tendency measures (i.e., comparison of means) in detecting changes in cognitive function in mTBI. Additionally, the findings highlight the utility of ANAM4's repeating a subtest at two points in a battery to explore within-subject differences in performance. (*JINS*, 2017, *23*, 1–6)

Keywords: Concussion, Military, Computerized assessment, Cognitive function, Neurocognitive assessment tool, Response consistency

INTRODUCTION

Every year thousands of service members (SMs) in the U.S. military are diagnosed with a mild traumatic brain injury (mTBI), also known as concussion (Defense and Veterans Brain Injury Center [DVBIC], 2016). These injuries can take place in a variety of settings due to several causes, including those similar to sports-related concussion in the civilian sector. Regardless of where or how concussion occurs, there is a need for timely and effective evaluation of an individual's cognitive functioning (e.g., Kelly, Coldren, Parish, Dretsch, & Russell, 2012). Assessment of cognitive abilities *via* neuropsychological (NP) tests is considered the cornerstone of concussion management (McCrory et al., 2013). However, these tests are time consuming and require particular expertise for administration and interpretation of results. In more recent years computerized neurocognitive assessment tools (NCATs) have been increasingly used as a quicker and more feasibly administered alternative to NP tests (e.g., Friedl et al., 2007; McCrory et al., 2013).

The Automated Neuropsychological Assessment Metrics 4 TBI-MIL (ANAM4) is an NCAT developed by the U.S. Army (Friedl et al., 2007) and widely used in the military (Defense Health Board, 2016). ANAM4 is regularly administered before a deployment as a means to generate a neurocognitive baseline for post-deployment and post-injury

Correspondence and reprint requests to: Wesley R. Cole, Intrepid Spirit, Womack Army Medical Center, Fort Bragg, NC 28310. E-mail: wesley.r.cole.ctr@mail.mil

comparison (DoDi 6490.13). Despite the goal of NCATs, including ANAM4, existing evidence is inconclusive regarding the ability to identify cognitive issues following concussion (see Arrieux, Cole, & Ahrens, 2017; Resch, McCrea & Cullum, 2013).

Typically findings from ANAM4 are based on analyses comparing post-injury scores either to individual baseline measurements or normative databases (see Haran, Dretsch, et al., 2016; McCrea et al., 2008). The most commonly interpreted standardized ANAM4 score is "throughput," calculated for each subtest and based on the number of correct responses per response time. Building on evidence from a growing literature, this article applies an alternative, more fine-grained scoring method that may be better suited for identifying cognitive dysfunction. The analyses focus on within-person inconsistent performance, or intra-individual neurocognitive variability.

Although intra-individual variability is often viewed as noise or test error, it may in fact reflect fluctuation in cognitive processing and reveal cognitive deficits that a mean or standard score is attempting, but failing, to capture. For example, research in aging populations has shown intra-individual variability on various behavioral and neurophysiological measures to be associated with decline in cognitive performance (for example, Fjell, Rosquist, & Walhovd, 2009; Lovden, Shing, & Linderberger, 2007). Although the literature base is relatively small, intra-individual variability in acute and post-acute concussion populations has been studied for more than 2 decades using both traditional NP and reaction time (RT) tests (e.g., Rabinowitz & Arnett, 2013; Sosnoff et al., 2007; Stuss et al., 1989).

Using NP tests, Hill, Rohling, Boettcher, and Meyers (2013) analyzed intra-individual variability using means from the Meyers Neuropsychological Battery in individuals reporting a history of mTBI and found that overall performance is negatively correlated with variability. Similarly, in a study using RT-based stimulus discrimination and flanker tests, history of concussion was shown to be associated with increased intra-individual variability (Parks et al., 2015). Beyond behavioral measures, Segalowitz, Dywan, and Unsal (1997) demonstrated for a TBI group, and not for a control group, RT variability was related to electrophysiological measures of attentional allocation and sustainment (the P300 amplitude and the preresponse component of the contingent negative variation E-Wave), supporting the idea that RT variability reflects this attentional processing.

Studies have also examined intraindividual-variability in TBI using NCATs. Bleiberg, Garmoe, Halpern, Reeves, and Nadler (1997) demonstrated participants with mild to moderate TBI performed more inconsistently in same-day and across multiple day sessions than a healthy control group. Makdissi et al. (2001) investigated a simple RT test in a different NCAT, CogState, in athletes and found greater standard deviation in reaction time in acutely concussed *versus* never concussed athletes at follow-up, though not at baseline. However, longer RT in concussed participants as compared to controls could account for greater standard deviation in RT. Sosnoff et al. (2007) adjusted for mean RT in a group of individuals tested within 72 hours of concussion and found that after this adjustment, concussed individuals did not have greater RT SD than healthy age- and gender-matched individuals.

The above studies, most of which demonstrate an ability to differentiate TBI and control group performance using intra-individual variability measures, all compare an individual's performance on a test or whole battery *across* test sessions. In contrast, the present investigation explores potential differences in intra-individual variability by comparing performance on one subtest repeated within a battery in patients with acute concussion and healthy controls. Our approach allows examination of the use of intra-individual variability analyses within an abbreviated window and without a need for repeat testing of an entire battery.

The ANAM4 is an ideal test to examine intra-individual variability in this way, as unlike most NCATs, the ANAM4 includes an identical simple RT (SRT) task at the beginning and the end of the battery. Although the ANAM4 standard output generates the RT standard deviation on each subtest, our approach differs because it examines the standard deviation of the difference between the trial-by-trial RT data. This approach allows for a more fine-grained measure of intra-individual variability and an individual's change in RT (i.e., dSRT) over a brief period of time. In addition to looking at trial-by-trial raw RT data and dSRT, the current study investigated acutely concussed individuals, as previous research suggests ANAM4 has limited clinical utility more than eight days following concussion, as well as healthy controls (e.g., Nelson et al., 2016). We hypothesize that this alternative trial-by-trial approach to interpreting RT on ANAM4 will reveal differences in variability and dSRT across the two groups.

METHODS

Sample

A total sample of 350 individuals was selected from a larger study's sample of SMs from Fort Bragg with and without mTBI where ANAM4 was administered (Cole, Arrieux, Dennison, & Ivins, 2017). Informed consent was obtained from all subjects and data were collected in compliance with the Womack Army Medical Center Institutional Review Board's regulations and requirements. The sample included 242 healthy controls (CTRL) and 108 participants within 7 days of mTBI (mTBI). The following criteria were used exclude data from the analyses: (1) potentially invalid data according to the ANAM4 embedded effort index (EI; CTRL: $n = 10$; mTBI: $n = 5$); and (2) RT less than 150 ms or greater than 900 ms (CTRL: $n = 1$; mTBI: $n = 3$), also deemed to be indicative of potentially invalid data.

After exclusions, 231 records were assigned to the CTRL group and 100 were assigned to the mTBI group.

https://doi.org/10.1017/S1355617717000856 Published online by Cambridge University Press

Instrumentation

The ANAM4 (CSRC, 2014) is an automated, computerized neurocognitive test battery that includes a sleepiness scale, mood scale, a self-report TBI questionnaire, and seven core subtests: Code Substitution Delayed (CDD), Code Substitution (CDS), Matching-to-Sample (M2S), Mathematical Processing (MTH), Procedural Reaction Time (PRO), Simple Reaction Time (SRT1), and Simple Reaction Time Repeated (SRT2). Due to the larger study's procedures, an additional battery of questionnaires was administered before testing, including demographics, military history, head injury history, Post-Traumatic Checklist – Civilian (PCL-C), and the Neurobehavioral Symptom Inventory (NSI). Following the questionnaires, the seven core ANAM subtests were administered per usual procedures. Validity of the data was evaluated by an embedded EI, which flags atypical scores based on accuracy and discrepancy of responses (Roebuck-Spencer, Vincent, Gilliland, Johnson, & Cooper, 2013). For the purposes of this manuscript, only the EI and the raw data from the SRT1 and SRT2 were used in the analyses.

Data Analyses

The following metrics were calculated using data from the SRT1 and SRT2 raw RT data ($N = 40$ trials): (1) SRT difference score (dSRT; formula 1), (2) the standard deviation (*SD*) of the dSRT (dSRT-*SD*), (3) the mean of dSRT, and (4) the standardized response mean (SRM) of dSRT (dSRT SRM; formula 2).

$$dSRT = SRT2 - SRT1 \quad (1)$$

$$SRM = \frac{dSRT}{sd} \quad (2)$$

Both dSRT SRM and dSRT-*SD* were used as metrics of intra-individual variability.

Statistical Analyses

Group differences for demographic data were examined using Mann-Whitney *U* tests and Chi-Square tests. There were minor violations of the Lilliefors test of normality for the simple reaction subtest data; however, the potential for a familywise type I error due to multiple comparisons was accounted for with sample sizes sufficient enough (i.e., $n > 30$) for the central-limit theorem to apply, robustness of the parametric tests used, and Bonferroni-Holm sequential corrections.

Group differences were analyzed using a general linear model (1×2) multivariate analysis of variance(MANOVA), with group membership (2 levels) as the between-subjects variable. Univariate tests and pairwise comparisons were conducted to follow-up significant main effects. Effect size (ES) for group differences was calculated using the Hedge's g and Cohen's U_3 statistic, and the results were interpreted using the following criteria: recommended minimum practical effect size (RMPE; ES > 0.41), moderate effect (ES > 1.15), and strong effect (ES > 2.70) (Ferguson, 2009).

All analyses were performed with Matlab 2015b (Mathworks, Natick, MA) and SPSS Version 22 (IBM, Armonk, NY).

RESULTS

There were significant differences for sex and rank on the demographic variables (Table 1). Differences in sex are believed to be due to the higher number of officers in the control group, as there was a higher proportion of female officers than female enlisted soldiers. It is believed that officers were over-represented in the control group due to their greater ability to control and dictate their daily schedules, allowing them to take time off to volunteer in a research study. There were no other statistically significant differences on other measured demographic variables (Table 1).

The results of the MANOVA revealed that there was a significant multivariate main effect for group membership on ANAM4 performance ($F_{(4,326)} = 18.56$; $p < .001$; $\eta_p^2 = .19$). The univariate tests associated with the main effect for group were significant for the SRT1 ($F_{(1,329)} = 23.93$; $p = .001$;, $\eta_p^2 = .07$), SRT2 ($F_{(1,329)} = 60.58$; $p < .001$; $\eta_p^2 = .16$), dSRT ($F_{(1,329)} = 14.77$; $p < .001$; $\eta_p^2 = .04$), dSRT-*SD* ($F_{(1,329)} = 55.22$; $p < .001$; $\eta_p^2 = .14$), and dSRT SRM ($F_{(1,329)} = 18.60$; $p < .001$; $\eta_p^2 = .05$). Pairwise comparisons revealed that the mean for the control group was significantly lower (i.e., faster RT, less variability, less change in RT over time) than the mean for the mTBI group on each metric (Table 2). It should be noted that the effect for group differences (ES > .41) exceeded the RMPE for each variable.

Table 1. Participant characteristics for control and mTBI groups

	CTRL (n = 231)	mTBI (n = 100)	p-Value
Characteristic			
Age, years			
Median	33	26	0.132[a]
Range	19–58	19–48	
Sex, male, *n* (%)	96 (97%)	187 (81%)	0.0001[b]
Rank, *n* (%)			
Enlisted	123 (53%)	90 (91%)	0.0003[b]
Officer	108 (47%)	9 (9%)	
Years active duty			
Median	10	3	0.101[a]
Range	0–28	1–24	
No. of deployments			
Median	2	1	0.380[a]
Range	0–11	0–7	
Education level[c]			
Median	4	3	0.361[a]
Range	1–5	1–5	

CTRL = control group; mild traumatic brain injury = mTBI group.
[a]Two-tailed Mann-Whitney U test.
[b]Chi-square test.
[c]0 = less than 12 years; 1 = General Educational Development (GED) certificate; 2 = high school graduate; 3 = some college; 4 = associate degree; 5 = bachelor's degree or higher.

https://doi.org/10.1017/S1355617717000856 Published online by Cambridge University Press

Table 2. Descriptive statistics for SRT subtest trial-by-trial raw data, for entire sample and enlisted sample only.

	CTRL			mTBI			Statistics			
Entire sample	*N*	*M*	*SD*	*N*	*M*	*SD*	Δ	tstat	ES	U_3
SRT1	231	292.66	46.98	100	329.21	88.44	−36.55	−4.89	−0.47	0.68
SRT2	231	279.38	34.88	100	333.37	91.28	−53.99	−7.78	−0.68	0.80
dSRT	231	−13.28	29.52	100	4.16	42.46	−17.44	−3.84	−0.45	0.60
dSRT sd	231	87.00	36.68	100	130.91	70.50	−43.91	−7.43	−0.70	0.83
dSRT SRM	231	−0.18	0.34	100	0.00	0.35	−0.18	−4.31	−0.52	0.62
Enlisted sample										
SRT1	123	292.97	47.49	91	329.91	87.77	−36.94	−3.96	−0.50	0.40
SRT2	123	281.04	35.27	91	335.29	91.21	−54.25	−6.02	−0.74	0.27
dSRT	123	−11.08	30.78	91	5.39	52.76	−16.473	−3.00	−0.37	0.37
dSRT sd	123	90.35	38.65	91	130.66	70.58	−40.31	−5.35	−0.68	0.33
dSRT SRM	123	−0.15	0.32	91	0.00	0.36	−0.15	−3.24	−0.43	0.37

CTRL = control group; mTBI = mild traumatic brain injury group; SRT = simple reaction time; SR2 = simple reaction time repeated; dSRT = SR2-SRT; dSRT sd = dSRT standard deviation; dSRT SRM = dSRT standardized response mean; Δ = CTL value − mTBI value; tstat = *t*-statistic; ES = effect size calculated as Hedge's *g*; U_3 = Cohen's U_3 statistic.

As a result of the group differences between enlisted and officer data, a second MANOVA was performed on just the data from enlisted personnel, which resulted in group equivalency across demographics. The results of the MANOVA revealed that there was a significant multivariate main effect for group membership on ANAM4 performance ($F_{(4,209)} = 10.12$; $p < .001$; $\eta_p^2 = .97$). The univariate tests associated with the main effect for group were significant for the SRT1 ($F_{(1,212)} = 15.71$; $p = .001$; $\eta_p^2 = .07$), SRT2 ($F_{(1,212)} = 36.25$; $p < .001$; $\eta_p^2 = .15$), dSRT($F_{(1,212)} = 8.98$; $p = .003$; $\eta_p^2 = .04$), dSRT-*SD* ($F_{(1,212)} = 28.57$; $p < .001$; $\eta_p^2 = .12$), and dSRT SRM ($F_{(1,212)} = 10.52$; $p < .001$; $\eta_p^2 = .05$). Pairwise comparisons revealed that the mean for the control group was significantly lower than the mean for the mTBI group for each metric with effects exceeding the RMPE for all variables except for dSRT (ES = −.37) (Table 2).

DISCUSSION

The current study investigated differences in mean RT and RT variability between healthy controls and those with acute concussion using raw trial-by-trial RT data from the ANAM4. This approach was relatively unique as most previous studies have focused on the use of standardized scores and cognitive efficiency metrics (e.g., throughput scores) to investigate group differences. Moreover, prior studies examining differences in RT variability have almost exclusively done so across test sessions rather than using a repeated subtest within a battery and test session. Our hypotheses were largely supported, as those with acute concussion had slower RTs and greater RT variability than healthy controls. The most important finding was that significant group differences were seen across all variables, with raw SRT2 and dSRT-*SD* appearing to be the most sensitive variables with ESs (−.68 and −.70, respectively) that were similar to values previously reported for raw SRT2 (ES = −.60; Adam et al., 2015) and nearly double values reported for throughput scores (ES = −.35; Haran, Alphonso, et al., 2016).

It is not surprising that there were differences in the variability between the healthy control and mTBI groups. RT and RT variability have been shown to provide information about the allocation of attentional resources in those with neurological insult such as mTBI. Specifically, it is thought that attention allocation can be measured by RT latency in healthy controls, whereas in those with mTBI attention allocation is more related to RT variability than RT latency (Bleiberg et al., 1997; Segalowitz et al., 1997). As such, the current finding of within subject variability on ANAM4 SRT performance in an acute mTBI group provides additional evidence to the body of literature.

In general, these results reveal greater trial-to-trial fluctuations in performance for the mTBI group as compared to the control group. Based on the central tendency theory, these fluctuations are often viewed as noise, instability, or error. However, they may be indicative of subtle cognitive decline after concussion that may otherwise be missed by more traditional metrics. That is, analyses of raw RT (particularly the raw SRT2 data), trial-by-trial RT change, and trial-by-trial RT variability appears to be an alternative metric for NCATs. Moreover, these alternative metrics may offer greater clinical utility than metrics commonly used in cognitive testing. Given the computerized nature of NCATs, metrics such as raw RT, trial-by-trial RT change, and trial-by-trial RT variability can be more quickly and feasibly calculated. Furthermore, ANAM4 presents a conceivable advantage over other NCATs by including a repeated simple reaction time test, allowing comparison of RT and RT variables across time though still within one testing session, potentially tapping into "cognitive fatigue."

Limitations

The current study was derived from data from a larger study, and, therefore, procedures not relevant to the current analyses surrounded this study's data collection of interest. These procedures sometimes included other testing before taking

https://doi.org/10.1017/S1355617717000856 Published online by Cambridge University Press

the ANAM4, which could have increased fatigue. However, any potential fatigue would be relatively equitable across groups and relatively controlled for by comparing SRT2 to SRT1 which occurred within the same testing session. Additionally, recent studies demonstrated that when administering multiple NCATs in one session, performance was not affected by the order of administration (Cole, Arrieux, Dennison, & Ivins, 2017; Nelson et al., 2016).

Another potential limitation is the differences between sex and rank in the control and mTBI groups, with more females in officers and more officers in control group. Even so, when omitting officers from analyses, thus rendering the groups equitable across all measured demographics, the results still held.

Finally, as with any study of NCATs, there are many factors that exist and were not controlled for. The computer platform used (e.g., hardware and software configurations), the participants' familiarity with the ANAM4, the nature of injury, time since injury (e.g., <3 days *vs*. 3–7 days), ongoing symptomatology, potential medication with cognitive side effects (e.g., stimulants or sedatives), and so on. However, all efforts were taken to administer the tests with a platform as close to the ANAM4 manual specifications. Additionally, testing was done in a quiet room with a trained test proctor, in an environment similar to how baseline or post-injury testing would likely occur, likely rendering the results ecologically valid despite the potential for other sources of error.

CONCLUSIONS AND FUTURE DIRECTIONS

The results from this study support a small but growing body of literature that raw RT, RT change, and RT variability scores may be much more sensitive to the subtle cognitive effects often seen after concussion. It appears that mTBI participants can temporarily perform similarly to normal controls on RT latency, but repeated RT assessments at multiple time points throughout a battery demonstrate increased inconsistent performance. Interpreting these metrics rather than the traditionally reported standardized scores (e.g., throughput) appears to hold promise for the use of ANAM4 in acute concussion populations.

In addition, this study highlights the strength of using raw scores instead of standardized scores, where subtle cognitive effects may be washed out. However, additional work is needed to fully clarify the clinical utility (e.g., diagnostic and prognostic capabilities) of these metrics, and to determine if they do indeed offer advantages over traditional metrics obtained from traditional NP tests and NCATs. There is some existing evidence that shorter ANAM4 SRT is predictive of recovery in those acutely concussed (Norris, Carr, Herzig, Labrie, & Sams, 2013). Thus, it may be that faster raw RT, less change in RT across SRT1 and SRT2, and less RT variability could be predictive of faster and/ or better recovery after concussion and, therefore, incorporated into return to duty or return to play decisions. Given the Army's baseline/ predeployment testing program, it will also be important to determine if baseline assessments are valuable with regard to such metrics for diagnostic and prognostic purposes.

ACKNOWLEDGMENTS

The authors thank Karen Schwab, Brian Ivins, Felicia Qashu, Mary Alice Dale, James Wes McGee, Katie Toll, and Alex Fender for their contributions to this research. This material is published by permission of the Defense and Veterans Brain Injury Center, operated by General Dynamics Information Technology for the U.S. Defense Health Agency under Contract No.W91YTZ-13-C-0015. The authors have no financial interests to disclose. The views expressed in this manuscript are those of the author(s) and do not necessarily reflect the official policy or position of the Department of the Navy, Department of the Army, Department of Defense, or the U.S. Government.

REFERENCES

Adam, O., Mac Donald, C.L., Rivet, D., Ritter, J., May, T., Barefield, M., … Brody, D.L. (2015). Clinical and imaging assessment of acute combat mild traumatic brain injury in Afghanistan. *Neurology*, *85*(3), 219–227.

Arrieux, J.P., Cole, W.R., & Ahrens, A.P. (2017). A review of the validity of computerized neurocognitive assessment tools in mild traumatic brain injury assessment. *Concussion*, *2*.

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. *Cognitive and Behavioral Neurology*, *10*(4), 247–253.

Cognitive Science Research Center, CSRC. (2014). *ANAM military battery, administration manual*. Norman, OK: CSRC.

Cole, W.R., Arrieux, J.P., Dennison, E.M., & Ivins, B.J. (2017). The impact of administration order in studies of computerized neurocognitive assessment tools (NCATs). *Journal of Clinical and Experimental Neuropsychology*, *39*(1), 35–45.

Defense Health Board. (2016). *Review of the scientific evidence of using population normative values for post-concussive computerized neurocognitive assessments*. Report prepared for the Undersecretary of Personnel and Readiness, Falls Church, VA: DHB.

Department of Defense Instruction (DoDI) 6490.13. (2013). *Comprehensive policy on traumatic brain injury-related neurocogitive assessments by the military services*. Washington, DC: Department of Defense.

Defense and Veterans Brain Injury Center. (2016). DoD worldwide numbers for TBI worldwide totals. Retrieved from http://dvbic.dcoe.mil/files/tbi-numbers/DoD-TBI-Worldwide-Totals_2000-2016_Q1_May-16-2016_v1.0_2016-06-24.pdf

Fjell, A.M., Rosquist, H., & Walhovd, K.B. (2009). Instability in the latency of P3a/P3b brain potentials and cognitive function in aging. *Neurobiology of Aging*, *30*(12), 2065–2079.

Friedl, K.E., Grate, S.J., Proctor, S.P., Ness, J.W., Lukey, B.J., & Kane, R.L. (2007). Army research needs for automated neuropsychologial tests: Monitoring soldier health and performance status. *Archives of Clinical Neuropsychology*, *22*, 7–14.

Ferguson, C.J. (2009). An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, *40*(5), 532–538.

Haran, F.J., Alphonso, A.L., Creason, A., Campbell, J.S., Johnson, D., Young, E., … Tsao, J.W. (2016). Reliable change estimates for assessing recovery from concussion using the ANAM4 TBI-MIL. *The Journal of Head Trauma Rehabilitation*, *31*(5), 329–338.

https://doi.org/10.1017/S1355617717000856 Published online by Cambridge University Press

Haran, F.J., Dretsch, M.N., Slaboda, J.C., Johnson, D.E., Adam, O.R., & Tsao, J.W. (2016). Comparison of baseline-referenced versus norm-referenced analytical approaches for in-theatre assessment of mild traumatic brain injury neurocognitive impairment. *Brain Injury*, *30*(3), 280–286. doi: 10.3109/02699052.2015.1118766

Hill, B.D., Rohling, M.L., Boettcher, A.C., & Meyers, J.E. (2013). Cognitive intra-individual variability has a positive association with traumatic brain injury severity and suboptimal effort. *Archives of Clinical Neuropsychology*, *28*(7), 640–648.

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.

Lovden, M., Li, S.C., Shing, Y.L., & Lindenberger, U. (2007). Within-person trial-to-trial variability precedes and predicts cognitive decline in old and very old age: Longitudinal data from the Berlin Aging Study. *Neuropsychologia*, *45*, 2827–2838.

Makdissi, M., Collie, A., Maruff, P., Darby, D.G., Bush, A., McCrory, P., & Bennell, K. (2001). Computerised cognitive assessment of concussed Australian Rules footballers. *British Journal of Sports Medicine*, *35*(5), 354–360.

McCrea, M., Pliskin, N., Barth, J., Cox, D., Fink, J., French, L., … Powell, M. (2008). Official position of the Military TBI Task Force on the Role of Neuropsychology and Rehabilitation Psychology in the Evaluation, Management, and Research of Military Veterans with Traumatic Brain Injury: APPROVED by: American Academy of Clinical Neuropsychology (AACN) American Psychological Association Division 40 (Neuropsychology) American Psychological Association Division 22 (Rehabilitation Psychology) National Academy of Neuropsychology (NAN). *The Clinical Neuropsychologist*, *22*(1), 10–26.

McCrory, P., Meeuwisse, W.H., Aubry, M., Cantu, B., Dvořák, J., Echemendia, R.J., … Sills, A. (2013). Consensus statement on concussion in sport: The 4th International Conference on Concussion in Sport held in Zurich, November 2012. *British Journal of Sports Medicine*, *47*(5), 250–258.

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.

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.

Parks, A.C., Moore, R.D., Wu, C.T., Broglio, S.P., Covassin, T., Hillman, C.H., & Pontifex, M.B. (2015). The association between a history of concussion and variability in behavioral and neuroelectric indices of cognition. *International Journal of Psychophysiology*, *98*(3), 426–434.

Rabinowitz, A.R., & Arnett, P.A. (2013). Intraindividual cognitive variability before and after sports-related concussion. *Neuropsychology*, *27*(4), 481–490.

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.

Roebuck-Spencer, T.M., Vincent, A.S., Gilliland, K., Johnson, D. R., & Cooper, D.B. (2013). Initial clinical validation of an embedded performance validity measure within the automated neuropsychological metrics (ANAM). *Archives of Clinical Neuropsychology*, *28*(7), 700–710.

Segalowitz, S.J., Dywan, J., & Unsal, A. (1997). Attentional factors in response time variability after traumatic brain injury: An ERP study. *Journal of the International Neuropsychological Society*, *3*(2), 95–107.

Sosnoff, J.J., Broglio, S.P., Hillman, C.H., & Ferrara, M.S. (2007). Concussion does not impact intraindividual response time variability. *Neuropsychology*, *21*(6), 796–802.

Stuss, D.T., Stethem, L.L., Hugenholtz, H., Picton, T., Pivik, J., & Richard, M.T. (1989). Reaction time after head injury: Fatigue, divided and focused attention, and consistency of performance. *Journal of Neurology, Neurosurgery, & Psychiatry*, *52*(6), 742–748.