Toward Return to Duty Decision-Making After Military Mild Traumatic Brain Injury: Preliminary Validation of the Charge of Quarters Duty Test

Mary V Radomski, PhD, OTR/L; Leslie F Davidson, PhD, OTR/L; Laurel Smith, MS, OTR/L; Marsha Finkelstein, MS; Amy Cecchini, PT, MS; Kristin J Heaton, PhD; Karen McCulloch, PT, PhD; Matthew Scherer, PT, PhD; Margaret M Weightman, PT, PhD

Abstract:

Introduction

Determining duty-readiness after mild traumatic brain injury (mTBI) remains a priority of the United States Department of Defense as warfighters in both deployed and non-deployed settings continue to sustain these injuries in relatively large numbers. Warfighters with mTBI may experience unresolved sensorimotor, emotional, cognitive sequelae including problems with executive functions, a category of higher order cognitive processes that enable people to regulate goal-directed behavior. Persistent mTBI sequelae interfere with warfighters' proficiency in performing military duties and signal the need for graded return to activity and possibly rehabilitative services. Although significant strides have been carried out in recent years to enhance the identification and management of mTBI in garrison (EXORD 165–13) and deployed settings (EXORD 242–11; DoDI 6,490.11), the Department of Defense still lacks reliable, valid, and clinically feasible functional assessments to help inform duty-readiness decisions. Traditional functional assessments lack face validity for warfighters and may have ceiling effects, especially as related to executive functions. Performance-based multitasking assessments have been shown to be sensitive to executive dysfunction after acquired brain injury but no multitasking assessments have been validated in adults with mTBI. Existing multitasking assessments are not ecologically valid relative to military contexts. A multidisciplinary military–civilian team of researchers developed and evaluated a performance-based assessment called the Assessment of Military Multitasking Performance. One of the Assessment of Military Multitasking Performance multitasks, the Charge of Quarters Duty Test (CQDT), was designed to challenge the divided attention, foresight, and planning dimensions of executive functions. Here, we report on the preliminary validation results of the CQDT.

Materials and Methods

The team conducted a measurement development study at Fort Bragg, NC, enrolling 83 service members (33 with mTBI and 50 healthy controls). Discriminant validity was evaluated by comparing differences in CQDT sub-scores of warfighters with mTBI and healthy controls. Associations between CQDT sub-scores and neurocognitive measures known to be sensitive to mTBI were examined to explore convergent validity. The study was approved by the Womack Army Medical Center Institutional Review Board (Fort Bragg).

Results

There were significant between-group differences in two of the four CQDT sub-scores (number of visits, $p = 0.012$; and performance accuracy, $p = 0.020$). Correlations between the CQDT sub-scores and some neurocognitive measures were statistically significant but weak, ranging from 0.287 (CQDT performance accuracy and NAB Numbers and Letters, Part D) to −0.421 (CQDT total number of visits and Automated Neuropsychological Assessment Metrics Tower Task). There were group differences in terms of participants' reading level, education, years in military, and stress symptoms; some of these characteristics may have influenced CQDT performance.

Conclusions

The CQDT demonstrated initial evidence of discriminant validity. Further study is warranted to more formally evaluate convergent/divergent validity and ultimately how and whether this performance-based multitasking measure can inform readiness to return to duty after mTBI.

Keywords:

Mild traumatic brain injury, post-concussion symptoms, executive function, outcomes assessment, military personnel

Determining duty-readiness after mild traumatic brain injury (mTBI) remains a concern of the United States Department of Defense[1–3] as warfighters in both deployed and non-deployed settings continue to suffer relatively large numbers of these injuries. Of the 17,672 traumatic brain injuries reported by the Department of Defense in 2016, 15,177 (85.9%) were classified as mild.[4] After mTBI, service members (SM) may experience impairments in executive functions,[5] especially with repeated injuries.[6] Executive functions are a category of higher order cognitive processes that together enable people to regulate goal-directed behavior.[7] Executive functions are central to the warfighter's success on the battlefield. For example, warfighters rely on intact executive functions to navigate in unfamiliar terrain (planning/organizing) while scanning the environment for the enemy (set-shifting between scanning and navigating task) and appropriately reacting in response to unanticipated people or events (initiation/inhibition). The same functions are often critical to success in managing personal and social roles.[8] Problems with executive functions and other unresolved sensorimotor, cognitive, and/or emotional sequelae of mTBI interfere with proficiency in performing military duties[9] and signal the need for graded return to activity and possibly rehabilitative services.[10] Medical decision-makers need information from reliable, valid functional assessments that integrate multi-modal sensitivity and ecological validity to help inform duty-readiness decisions.[9] Doing so is challenging as traditional functional assessments have ceiling effects[11] and lack face validity for SM with mTBI-related impairments.[12]

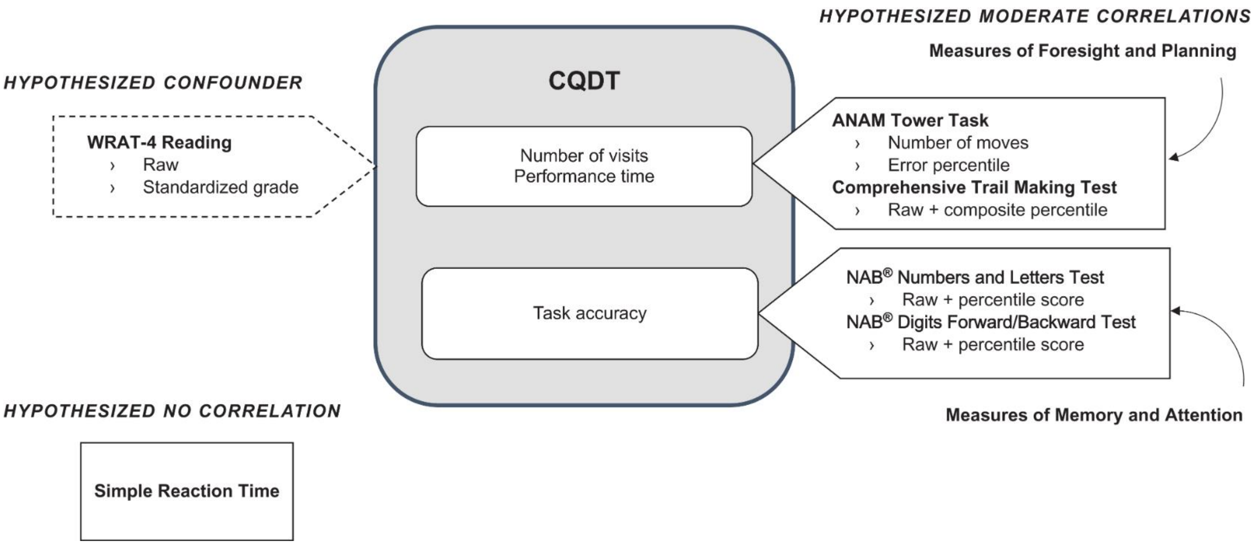
Performance-based assessment is a form of functional assessment composed of tasks with properties that simulate the demands of everyday life. Examiners use observational, behaviorally based metrics to characterize various aspects of the patient's capabilities.[13] Performance-based multitasking assessments were developed to be sensitive to executive dysfunction after traumatic brain injury and mild stroke.[14–16] Examiners assess patients' efficiency of goal processing and pursuits[17] by observing performance of complex activities organized within larger ecologically valid contexts or scenarios[18] such as shopping (Multiple Errands Test [MET])[19] or working in a library (Complex Task Performance Assessment [CTPA]).[20]Multitasking assessments typically include several common features: many tasks are required; tasks are dovetailed; only one task is performed at a time; interruptions occur unexpectedly; and demands are placed on prospective memory.[21] Because ecologically valid performance-based multitasking tests typically lack the structure and organization inherent in most traditional or laboratory-based tests,[22] they may better approximate real-life demands and predict the quality of everyday functioning.[23] At present, there are no multitasking assessments specifically validated on adults with mTBI. Additionally, none of the existing multitasking assessments are ecologically valid relative to military contexts.

Informed by this literature and end-user and expert input,[12] a military–civilian team of rehabilitation and psychology researchers developed a performance-based assessment called the Assessment of Military Multitasking Performance (AMMP). The AMMP battery is composed of three dual-task and three multitask tests that simulate sensorimotor, cognitive, and/or exertional demands of various soldiering tasks. It is designed to fulfill the need for a rigorous functional assessment to help inform duty-readiness decision-making after mTBI. One of the AMMP multitasks, the Charge of Quarters Duty Test (CQDT),[24] uses the structure of the MET[25] and Burgess' definition of multitasking[21] to challenge the divided attention, foresight, and planning dimensions of executive functions.

The research team conducted a proof-of-concept study to examine the inter-rater reliability and preliminary validity of the six AMMP subtests. (See a complete description of iterative inter-rater reliability testing in Weightman et al.[26]) Here, we report on the preliminary validation results of the CQDT. Discriminant/known-group validity was evaluated by comparing the CQDT sub-scores of two groups of SM: (1) SM who were not duty-ready due to ongoing mTBI symptomatology and (2) SM considered to be healthy and duty-ready. Convergent validity was explored by comparing participants' CQDT sub-scores with their scores on an array of neurocognitive measures. We hypothesized that CQDT sub-scores of SM with known mTBI would significantly differ from those who were healthy. Further, we hypothesized that there would be significant, weak-to-moderate correlations between CQDT sub-scores and scores on neurocognitive measures of foresight and planning

(Fig. 1). Confirmation of these hypotheses in this initial psychometric evaluation would provide evidence that the CQDT warrants further study toward predicting duty-readiness after mTBI.

Figure 1.



Hypothesized associations between CQDT sub-scores and neurocognitive measures.

Methods

This measurement development study was conducted at Fort Bragg, NC. The study received approval from the Womack Army Medical Center Institutional Review Board at Fort Bragg. Written informed consent was obtained from each participant.

Participants

A convenience sample of two groups of active duty SM at Fort Bragg (18–42 yr of age) was recruited. Participants designated as healthy controls were performing military duties without restrictions, eligible for deployment, and denied any type of concussion or head injury within the preceding 24 mo or had no residual symptoms from a remote concussion. These individuals were recruited from tenant units on Fort Bragg via informal unit and Defense Veterans Brain Injury Center briefings and postings placed around the base. Participants in the mTBI comparative group had not been cleared for full duty due to ongoing mTBI symptoms. These individuals had an mTBI documented in the medical record and were receiving rehabilitation services for mTBI sequelae at the Womack Army Medical Center TBI Clinic. Participants either self-referred to the study based on installation briefings and recruitment flyers (with subsequent screening from study personnel) or were referred by rehabilitation personnel based on meeting the following inclusion criteria: mTBI occurring from 2 wk to 2 yr before the AMMP test date; sufficient vision (corrected or uncorrected) for reading and performance of everyday tasks; able to hear close-range or telephone conversation without amplification; sufficient stamina to perform everyday activities that require moderate exertion (12–14 on the Borg Rating of Perceived Exertion);[27] and tolerate a 2- to 3-h testing session with frequent breaks. SM in either group were excluded if they had (1) documented duty-limiting physical or behavioral health conditions that prevented moderate exertion for 10 min or continuous activity for up to 30 min or (2) a history of psychiatric disorder, moderate or severe brain injury, or penetrating head injury.

Data Collection

Subjects participated in a one-time testing session lasting up to 3 h. Physical and occupational therapists and a research assistant collected study data. Subjects completed an intake questionnaire, after which study personnel administered neurocognitive tests. AMMP subtests were then administered by a physical or occupational therapist in varying order in an effort to minimize bias from order effects.

Sociodemographic and Symptom Characteristics

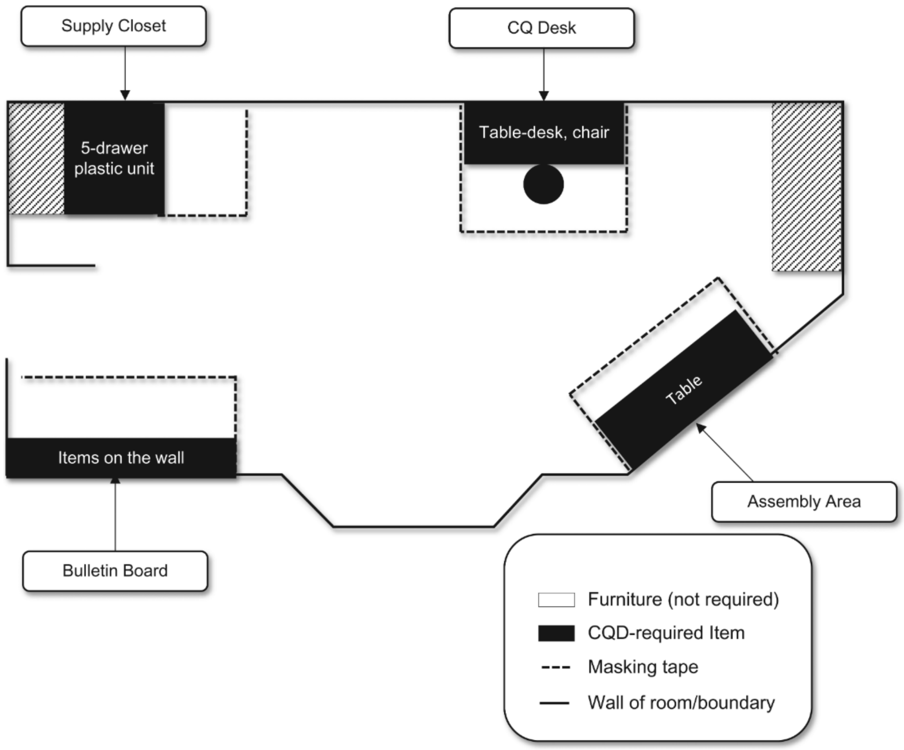
Participants provided the following information via the intake and other questionnaires: age, ethnicity, education level, first language, and learning disabilities; military history (pay grade, length in military service, current military occupational specialty, number of deployments, and duration of deployment); current pain and energy level; injury and behavioral health history; recent sleep history; hearing impairments; mTBI-related symptoms (Neurobehavioral Symptom Inventory);[28] stress-related symptoms (Post Traumatic Stress Disorder Checklist-Civilian);[29] and a question about perceived readiness to be deployed to a combat zone in 72 h. The Test of Memory Malingering (TOMM),[30] a test of exaggerated or deliberately faked memory impairment, was administered to verify performance effort. The Wide Range Achievement Test

Version 4 Reading Test[31] was administered as an estimate of educational background and intelligence.

Charge of Quarters Duty Test

The CQDT is a multitasking assessment that was developed through an iterative process using the structure and metrics of the MET,[25] a measure with known sensitivity to executive dysfunction.[32] In the CQDT, the SM must complete a list of assignments related to a CQ shift, a common military duty. Assignments include assembling a PVC footstool (Formufit. 15342 S. Keeler St., Suite B, Olathe, KS 66062), reporting information to a supervisor at various times, and inventorying supplies.[24]Participants are also told to adhere to task rules and complete the assignments making as few visits as possible to the four work stations, designated with signage and masking taped boundaries (Fig. 2). There are four performance sub-scores: number of visits (number of times one of the four work stations was visited during the exercise, six visits representing the best possible efficiency), performance accuracy (number of points for task completion, 38 total), performance time (in minutes), and total number of times that the subject broke task rules. Inter-rater reliability of CQDT sub-scores were clinically acceptable with demonstrated intraclass correlations ranging from 0.88 to 1.0.[26] Although all participants were likely familiar with the task context (performing charge of quarters duty), the CQDT represented a novel task because of the specific assignment requirements and the challenge to complete the assignments in as few visits to work stations as possible.

Figure 2.



Four CQDT work areas can be flexibly set up in most clinical and testing spaces.

Neurocognitive Measures

Neurocognitive measures were chosen that had known sensitivity to cognitive vulnerabilities associated with mTBI, could be administered by study personnel, and were sufficiently brief to avoid unreasonable test burden. For participants with mTBI, scores on the Neuropsychological Assessment Battery (NAB),[33]Comprehensive Trail-Making Test (CTMT),[34] and TOMM were obtained from the Womack Army Medical Center medical record. These tests were administered as part of the TBI Pipeline Assessment for the current episode of care a few weeks to a few months before AMMP testing. Study personnel administered these tests to healthy control SM and the other tests listed below to all participants during the testing session (see Table I).

Table I.

Neurocognitive Measures Administered to Evaluate Convergent Construct Validity

Neurocognitive Measure	Cognitive Domain	Description
Simple Reaction Time (SRT) from the *Automated Neuropsychological Assessment Metrics, Version 4* (ANAM4) (Center for the Study of Operator Performance, 2007)	Reaction time	The participant clicks the mouse as rapidly as possible when presented with a series of symbols on a computer display. It was administered at the beginning of all testing and at completion of all testing as a probe of fatigue
The *ANAM4 Tower Test* (Center for the Study of Operator Performance, 2007)	Executive functions: foresight and planning	Based on the *Tower of Hanoi* (Welsh, 2001), the participant figures out how to move a stack of different-sized disks from one rod to one of the other two rods while adhering to three rules; administered via computer
CTMT (Reynolds, 2002)	Attention, visual scanning, resistance to distraction, set-shifting	The participant connects a series of stimuli in a specified order as fast as possible across five trials. Distractor stimuli on some trials increase the demands of inhibitory control
NAB Numbers and Letters Test	Attention, psychomotor speed, processing speed, impulse control	The participant performs four subtests that incorporate letter cancellation and counting
NAB Digits Forward/Digits Backward Test	Working memory and attention	The measure has both a seven-item digits forward task and a seven-item digits backward task, each with its own individual score

Statistical Analyses

Data were entered into a RedCap[35] database and data analysis was conducted with SPSS Version 22 (IBM SPSS Inc., Chicago, IL). To preserve as many subjects as possible and avoid introduction of additional bias, we followed the principle of "Full Analysis Set," including all participants that started the CQDT regardless of their performance. Descriptive analyses were used to examine participants' sociodemographic and symptom characteristics. Using cutoff scores proposed by Jones,[36] individuals whose TOMM scores suggested poor effort were excluded from all analyses.

Discriminant/known-group validity was evaluated by comparing CQDT sub-scores between healthy controls and those with mTBI. A sample size of 50 and 30 in each group, healthy controls and SM with mTBI, respectively, provided 80% power to detect an effect size of 0.3 (medium effect size for non-parametric statistic) at a two-sided alpha of 0.05. Because scores were skewed, the non-parametric Mann–Whitney *U* test was used to evaluate between-group differences where a two-sided p-value <0.05 was required for significance. Non-parametric effect size was calculated as the Z-score from the Mann–Whitney *U* divided by the total observations.

Convergent validity was assessed using a Pearson's correlation coefficient that included CQDT sub-scores and neurocognitive domain scores. Where possible, raw unadjusted neurocognitive scores were used to parallel unadjusted CQDT scores. The sample size of 80 provided 80% power to detect a correlation for expected convergence at a minimum of 0.30 at a two-sided alpha of 0.05. Weak–moderate correlation (i.e., r between 0.3 and 0.5) between CQDT sub-scores and neurocognitive domains was expected based on convergent validity of other performance-based multitasking assessments.[20,37]

Results

Descriptive Analyses

In total, 84 SM from Fort Bragg agreed to participate in the study, but one healthy control was excluded from the analysis based on the evidence of insufficient effort on the TOMM. Table II lists sociodemographic and symptom characteristics of 83 individuals whose data were analyzed in this study. Those with mTBI were significantly younger, had fewer years of education and military service, lower reading levels, and more significant stress symptoms than healthy control participants.

Table II.

Characteristics of Participants Performing the CQDT

Characteristic	Healthy Controls, *N* = 50	mTBI, *N* = 33	*p*-Value
Age (yr)	30.2 ± 6.1	26.2 ± 5.2	0.001[a]
Sex			
Women	10 (20.0)	2 (6.1)	0.112[b]
Men	40 (80.0)	31 (93.9)	
Race/ethnicity			
Non-Hispanic White	25 (50)	21 (63.6)	
Non-Hispanic Black	15 (30.0)	4 (12.1)	0.273[b]
Hispanic	6 (12.0)	3 (9.1)	
Asian	3 (6.0)	3 (9.1)	
Other	1 (2.0)	2 (6.1)	
Education			
High school	6 (12.0)	6 (18.2)	0.008[b]
Trade school	1 (2.0)	2 (6.1)	
Some college	20 (40.0)	22 (66.7)	

Characteristic	Healthy Controls, *N* = 50	mTBI, *N* = 33	*p*-Value
Bachelor's degree	17 (34.0)	3 (9.1)	
Advanced degree	6 (12.0)	0 (0.0)	
Years in military (*n* = 49, 31)	8.4 ± 5.5	5.2 ± 4.6	0.004[a]
Reading level: WRAT-4 (raw reading)	61.1 ± 5.5	58.1 ± 6.0	0.018[a]
Stress symptoms:	22.2 ± 8.2	34 ± 14.7	<0.001[c]
PCL-C sum (*n* = 50, 31)	19 (17−63)	32 (17–73)	

PCL-C, Post Traumatic Stress Disorder Checklist-Civilian; WRAT-4, Wide Range Achievement Test Version 4.

NOTE. Values are *n* (%), mean ± standard deviation. Median (range) included for PCL-C sum.

[a]*t*-Test.

[b]Chi-square.

[c]Mann–Whitney *U*.

Discriminant/Known-Group Validity

Healthy control SM completed the CQDT in fewer visits ($p = 0.012$) and with greater accuracy ($p = 0.020$) than SM with mTBI (Table III). Correlation between visits and accuracy was low ($r = -0.31$), indicating that each of these performance measures can provide information for evaluating recovery from mTBI. Group differences did not reach significance for performance time or number of times the rules were broken.

Table III.

CQDT Scores

Scoring Item	Healthy Controls, *N* = 50	mTBI, *N* = 33	p-Value* (Effect Size)[a]
Number of visits	12.7 ± 6.7	14.8 ± 4.9	0.012 (0.27)
	11.5(6–51)	14(8–29)	
Performance accuracy	34.4 ± 2.8	32.6 ± 3.9	0.020 (0.25)
	35(29–38)	34(21–37)	
Total performance time (min)	19.0 ± 4.8	20.8 ± 5.5	0.171
	18.1(11.7–31.9)	19.6(13.1–37.0)	
Total number of times that rules were broken	1.1 ± 2.5	1.3 ± 2.0	0.512
	0 (0–13)	0 (0–6)	

NOTE. Values are mean ± standard deviation, median (range).

*Mann–Whitney *U* test for continuous variables and Chi-square for categorical variable.

[a]Effect size provided for significant elements.

Convergent Validity

Between-group comparisons on neurocognitive measures suggested statistically significant differences in CTMT, two of the NAB Numbers and Letters tests, and NAB Digit Forward–Backward (Table IV). Additionally, there were between-group differences in post-test simple reaction time, a metric to characterize influence of fatigue. Overall correlations between the CQDT sub-scores and some neurocognitive measures were statistically significant but weak, ranging from 0.287 (CQDT performance accuracy and NAB Numbers and Letters, Part D) to −0.421 (CQDT total number of visits and Automated Neuropsychological Assessment Metrics [ANAM] Tower Task) (Table V). CQDT performance accuracy was significantly correlated (weak) with simple reaction time scores at baseline and end of test session (−0.328 and −0.406, respectively). CQDT sub-scores were not correlated with NAB Digits Forward–Backward scores.

Table IV.

Comparison Between Subject Groups on Neurocognitive Measures

Test (# of HC, # of mTBI)		**HC**	**mTBI**	***p*-Value**
Simple reaction time	Baseline	279.4 ± 39.0	312.3 ± 97.1	0.372
		274.9 (223.7–453.5)	284.8 (208.3–647.7)	
	End of testing	258.0 ± 41.0	343.8 ± 160.9	0.003
		254.3 (214.2–509.5)	279.3 (213.2–970.6)	
ANAM4 Tower Test (50, 33)	Number of moves	9.9 ± 2.1	10.6 ± 2.5,	0.167
		9.7 (7.2–15.4)	10 (7.2–19.6)	
CTMT (50, 26)	Composite index	46.5 ± 9.6	38.6 ± 9.6	0.022*
		47.0 (29.0–68.0)	39.5 (17.0–62.0)	
NAB Numbers & Letters (50, 27)	Part A raw	98.4 ± 18.7	90.9 ± 18.7	0.047*
		97.5 (62.0–141.0)	87.5 (63.0–141.0)	
	Part B raw	96.9 ± 21.5	75.8 ± 20.7	<0.001**
		97 (31.0–143.0)	81.0 (27.0–108.0)	
	Part C raw	58.6 ± 17.8	50.4 ± 24.5	0.186

Test (# of HC, # of mTBI)		HC	mTBI	*p*-Value
		57.5 (15.0–96.0)	50.0 (14.0–87.0)	
		49.1 ± 10.1	43.4 ± 10.9	
	Part D raw	47 (30.0–73.0)	45.5 (16.0–64.0)	0.062
		8.6 ± 2.4	8.3 ± 2.3	
	Forward raw	8.0 (3.0–14.0)	8.0 (4.0–13.0)	0.484
	Backward raw	5.5 ± 2.4	4.1 ± 1.7	
NAB Digit Forward–Backward (50–25)		5.0 (0.0–11.0)	4.0 (2.0–9.0)	0.019*

NOTE. Values are mean ± standard deviation, median (range).

*Significance at the 0.05 level (two-tailed).

**Significance at the 0.01 level (two-tailed).

Table V.

CQDT Sub-score Correlations with Neurocognitive Metrics (*r*, Pearson Correlation Coefficients)

		Performance accuracy	**Total Number of Times Rules were broken**	**Total Number of Visits**	**Performance Time**
ANAM4 Tower Test (*n* = 83)	Number of moves	−0.301**	0.246*	−0.421**	0.239*
	Part A raw	0.038	0.015	−0.022	−0.310**

		Performance accuracy	Total Number of Times Rules were broken	Total Number of Visits	Performance Time
NAB Numbers & Letters (*n* = 76)	Part B raw	0.218	−0.031	−0.155	−0.406**
	Part C raw	0.287*	−0.131	−0.240*	−0.195
	Part D raw	0.235*	−0.140	−0.242*	−0.262*
NAB digits forward–backward (*n* = 74)	Digits forward raw	0.073	−0.054	−0.077	−0.107
	Digits backward raw	0.191	−0.087	−0.121	−0.107
CTMT (*n* = 76)	Composite index	0.278*	−0.159	−0.225	−0.399**
Simple reaction time (*n* = 83)	Baseline	−0.328**	0.069	−0.040	−0.019
	End of testing	−0.406**	0.038	−0.019	−0.041

*Correlation is significant at the 0.05 level (two-tailed).

**Correlation is significant at the 0.01 level (two-tailed).

Discussion

The AMMP, a performance-based battery of dual- and multitask tests, was developed to help inform duty-readiness decisions for SM with mTBI. The CQDT is an AMMP subtest designed as a multitask to challenge executive functions. This study provides preliminary evidence that the CQDT has initial validity and warrants further investigation. Results of hypothesis testing supported known-group discriminant validity with significant, between-group differences in two of the four CQDT sub-scores (number of visits, $p = 0.012$; and performance accuracy, $p = 0.020$). Effect sizes (0.27 and 0.25, respectively) were small at less than 0.3.[38]

Hypotheses related to convergent construct validity were partially confirmed. The CQDT challenges participants to complete the assignments in as few visits as possible to the four work stations. Therefore, weak-to-moderate correlation between CQDT number of visits and neurocognitive metrics that challenged foresight and planning (ANAM Tower Task and CTMT) was hypothesized. In fact, there were statistically significant but weak correlations between CQDT sub-scores and neurocognitive measures that challenge foresight and planning (i.e., ANAM Tower Test and CTMT) but no correlation with measures of working memory (NAB Digits Forward–Backward). Correlations between the Wide Range Achievement Test Version 4 and CQDT sub-scores were also weak, suggesting that premorbid intelligence may not have been a strong confounder. Although participants with mTBI may have been more fatigued than healthy controls at post-test (as evidenced by between-group differences in simple reaction time), the association between CQDT performance accuracy and simple reaction time was significant but weak, suggesting that fatigue may not have had a strong influence on CQDT performance.

Preliminary CQDT validation findings are similar to that of the MET-Revised[39] and CTPA,[40] which like the CQDT, and were developed to place demands on executive function using the Burgess definition of multitasking.[21] Both have been validated on adults with mild stroke.[20,39] Similar to those with mTBI, patients with mild stroke are vulnerable to subtle but potentially disabling problems with executive functions that may go undetected by traditional rehabilitation assessments.[16,41] Also using known-group analyses, Morrison and colleagues found statistically significant group differences between adults with mild stroke and healthy age-matched controls on MET-R sub-scores for total tasks completed ($p = 0.001$), and number of rule breaks ($p = 0.001$).[39] Unlike the CQDT, the MET-R total locations (similar to CQDT total visits) only approached significance ($p = 0.08$); between-group differences in performance time were not significant for either the MET-R or CQDT. Wolf and colleagues also found statistically significant group differences between community controls and adults with mild stroke on sub-scores of the CTPA (total score, number of task failures, total number of inefficiencies, and total rule breaks).[20] There were weak correlations between the total CTPA score and the Weschler

Test of Adult Reading ($r = -0.49$, $p = 0.003$) and the Delis–Kaplan Executive Function System color–word combination test ($r = -0.43$, $p = 0.003$).[20]

Weak correlations between performance-based assessments of multitasking (e.g., CQDT and CTPA) and domain-specific neurocognitive measures should be anticipated as these two assessment approaches measure intersecting but distinct constructs. Neurocognitive measures attempt to characterize cognitive capacity in specific, isolated cognitive domains, whereas performance-based assessment characterizes complex performance by simultaneously placing demands a wide array of interacting and integrated cognitive and sensorimotor systems. Furthermore, neurocognitive data were collected at different time points based on group, which may confound interpretation of these data. For healthy controls, neurocognitive tests were administered during their AMMP testing session, whereas neurocognitive data were taken from medical records of SM with mTBI. Neurocognitive tests were administered to participants with mTBI during a period of weeks to months preceding their AMMP testing session. These data were retrospectively obtained from the medical record in order to minimize test burden for participants with mTBI, which may have inadvertently introduced confounds. This issue should be addressed in future studies along with conducting comparisons between CQDT scores and other performance-based multitasking assessments (i.e., MET or CTPA) or a relevant real-life parameter, neither of which was feasible within the context of the AMMP study.

Recognizing the limitations of sole reliance on self-report and assessments of isolated cognitive and sensorimotor domains, the AMMP battery was developed to address the Army's need for a multidimensional examination of a SM's readiness to return to potentially high-stakes military responsibilities after mTBI. Performance-based assessment may fill an existing gap in comprehensive assessment after mTBI because it characterizes what a person can *do* versus focusing on his or her impairments or complaints. This emphasis on performance may be of particular importance to post-mTBI assessment because of the difficulty directly correlating symptoms and impairments, given the multifactorial origins of persistent post-mTBI problems.[42] The AMMP is the only performance-based assessment for mTBI that incorporates a multitasking subtest designed specifically to challenge executive functions in the context of complex tasks. In addition to meeting clinical feasibility requirements (low cost and space requirements), the CQDT was constructed to offer minimal structure, a novel environment, and maximal challenge in the assumption that executive functions are only activated when a person perceives that an immediate task is important and cannot be managed with behavioral habits.[43]Because such circumstances are necessary to elucidate dysexecutive symptoms,[43]standardized, scorable, and logistically feasible performance-based multitasking assessments like the CQDT (and its civilian counterpart now in-development) may prove to be essential to decision-making and rehabilitation planning after mTBI in both military and civilian contexts in the future.

Limitations

Although preliminary validation findings are encouraging, the group differences in sociodemographic characteristics must be considered in evaluating study findings and planning future studies. There were statistically significant group differences in reading level (proxy for premorbid intelligence), years of education, and years in the military, which may have contributed to differences in CQDT performance. Higher reading levels, years of education, and military service suggest that there were likely more commissioned officers and senior enlisted personnel in the healthy control group than in the mTBI cohort. Further study can evaluate at least two possible explanations. First, it may be that junior enlisted personnel, who tend to be younger with less education, are disproportionately vulnerable to mTBI. Second, officers and senior enlisted personnel had greater schedule flexibility and motivation to volunteer as a research participant than younger, enlisted soldiers. Because this proof-of-concept study, we relied on a convenience sample of volunteers. As with other reports,[44] participants with mTBI in this study had more significant stress symptoms, which in and of themselves can contribute to executive dysfunction.[45]Finally, although known groups were used to test construct validity, the actual groups may not have been as “known” as hypothesized. The research team did not have specific information regarding the symptomatology for which the SM with mTBI were receiving rehabilitation services. Therefore, within the mTBI group, SM may not have had lingering cognitive or executive dysfunction but rather sensorimotor difficulties, unlikely to be challenged by the CQDT, possibly explaining smaller effect sizes.

Conclusions

In this study, the CQDT demonstrated initial evidence of validity, suggesting that it is worthy of further study. The CQDT represents a promising, potential addition to the rather limited number of psychometrically sound, clinically feasible, performance-based multitasking assessments available to rehabilitation clinicians seeking to identify individuals with executive dysfunction. Further study will determine whether contextually rich performance-based multitask tests like the CQDT represent a more sensitive alternative or supplement to neurocognitive testing when determining readiness to return to high-stakes activities within both civilian and military contexts following potentially concussive events.

Acknowledgments

The AMMP validation study was funded by the United States Army Medical Research Materiel Command (award number W81XWH-12–2-0070). The authors gratefully acknowledge the important contributions of the following individuals: CAPT Henry McMillan, Ms. Caroline Cleveland, Dr. Arthur Estrada, and Dr. Edward Zambraski. We acknowledge Emily Shoemaker, MPH of the Courage Kenny Research Center for her invaluable assistance with Figures 1 and 2.

Funding

United States Army Medical Research Materiel Command (award number W81XWH-12-2-0070).

References (Endnotes)

1 Department of the Army : Guidance for management of concussion/mild traumatic brain injury in the garrison setting. HQDA EXORD 165–132015, Washington, D.C. May 18, 2015. Available at www.qmo.amedd.army.mil/mTBI/ConcussionManagementGarrison.pdf; accessed September 22, 2017.

2 Department of Defense : Policy guidance for management of concussion/mild traumatic brain injury in the deployed setting. DoDI 6490.112012, Washington D.C. September 18, 2012. Available at http://www.usaisr.amedd.army.mil/cpgs/DODI_6490.11_Policy_Guidance_for_Mgmt_of_Mild_Traumatic_Brain_Injury_or_Concussion_in_the_Deployed_Setting.pdf; accessed September 22, 2017.

3 Helmick KM , Spells CA, Malik SZ, Davies CA, Marion DW, Hinds SR: Traumatic brain injury in the US military: epidemiology and key clinical and research programs. Brain Image Behav 2015; 9(3): 358–66.

4 Defense Veterans Brain Injury Center (DVBIC) : DoD numbers for traumatic brain injury: worldwide totals 2015. Available at http://dvbic.dcoe.mil/files/tbi-numbers/DoD-TBI-Worldwide-Totals_2015_May-16-2016_v1.0_2016-06-24.pdf. Accessed August 11, 2016.

5 Howell D , Osternig L, Van Donkelaar P, Mayr U, Chou LS: Effects of concussion on attention and executive function in adolescents. Med Sci Sports Exerc 2013; 45(6): 1030–7.

6 Seichepine DR , Stamm JM, Daneshvar DH, et al. : Profile of self-reported problems with executive functioning in college and professional football players. J Neurotrauma 2013; 30(14): 1299–1304.

7 Levine B , Robertson IH, Clare L, et al. : Rehabilitation of executive functioning: an experimental-clinical validation of goal management training. J Int Neuropsychol Soc 2000; 6(3): 299–312.

8 Chan RC , Shum D, Toulopoulou T, Chen EY: Assessment of executive functions: review of instruments and identification of critical issues. Arch Clin Neuropsychol 2008; 23(2): 201–16.

9 Scherer MR , Weightman MM, Radomski MV, Davidson LF, McCulloch KL: Returning service members to duty following mild traumatic brain injury: exploring the use of dual-task and multitask assessment methods. Phys Ther 2013; 93(9): 1254–67.

10 McCulloch KL , Goldman S, Lowe L, et al. : Development of clinical recommendations for progressive return to activity after military mild traumatic brain injury: guidance for rehabilitation providers. J Head Trauma Rehabil 2015; 30(1): 56–67.

11 Hall KM , Mann N, High W, et al. : Functional measures after traumatic brain injury: ceiling effects of FIM, FIM+FAM, DRS, and CIQ. J Head Trauma Rehabil 1996; 11: 27–39.

12 Radomski MV , Weightman MM, Davidson LF, et al. : Development of a measure to inform return-to-duty decision making after mild traumatic brain injury. Mil Med 2013; 178(3): 246–53.

13 Augustin A , Loewenstein A, Kuppermann BD: Macular edema. General pathophysiology. Dev Ophthalmol 2010; 47: 10–26.

14 Frisch S , Forstl S, Legler A, Schope S, Goebel H: The interleaving of actions in everyday life multitasking demands. J Neuropsychol 2012; 6: 257–69.

15 McCulloch KL , Mercer V, Giuliani C, Marshall S: Development of a clinical measure of dual-task performance in walking: reliability and preliminary validity of the Walking and Remembering Test. J Geriatr Phys Ther 2009; 32(1): 2–9.

16 Morrison MT , Edwards DF, Giles GM: Performance-based testing in mild stroke: identification of unmet opportunity for occupational therapy. Am J Occup Ther 2015; 69(1): 6901360010p1–5.

17 Toplak ME , West RF, Stanovich KE: Practitioner review: do performance-based measures and ratings of executive function assess the same construct?J Child Psychol Psychiatry 2013; 54(2): 131–43.

18 Burgess PW , Alderman N, Forbes C, et al. : The case for the development and use of “ecologically valid” measures of executive function in experimental and clinical neuropsychology. J Int Neuropsychol Soc 2006; 12(2): 194–209.

19 Alderman N , Burgess PW, Knight C, Henman C: Ecological validity of a simplified version of the multiple errands shopping test. J Int Neuropsychol Soc 2003; 9(1): 31–44.

20 Wolf TJ , Dahl A, Auen C, Doherty M: The reliability and validity of the Complex Task Performance Assessment: a performance-based assessment of executive function. Neuropsychol Rehabil 2017; 27(5): 707–721.

21 Burgess PW : Strategy application disorder: the role of the frontal lobes in human multitasking. Psychol Res 2000; 63(3–4): 279–88.

22 Salthouse TA , Atkinson TM, Berish DE: Executive functioning as a potential mediator of age-related cognitive decline in normal adults. J Exp Psychol Gen 2003; 132(4): 566–94.

23 Spooner DM , Pachana NA: Ecological validity in neuropsychological assessment: a case for greater consideration in research with neurologically intact populations. Arch Clin Neuropsychol 2006; 21(4): 327–37.

24 Smith LB , Radomski MV, Davidson LF, et al. : Development and preliminary reliability of a multitasking assessment for executive functioning after concussion. Am J Occup Ther 2014; 68(4): 439–43.

25 Shallice T , Burgess P: Deficits in strategy application following frontal lobe damage in man. Brain 1991; 114: 727–41.

26 Weightman MM , McCulloch KL, Radomski MV, et al. : Further development of the assessment of military multitasking performance: iterative reliability testing. PLoS One 2017; 12(1): e0169104.

27 Borg GA : Psychophysical bases of perceived exertion. Med Sci Sports Exerc 1982; 14(5): 377–81.

28 Cicerone KD , Kalmar K: Persistent postconcussion syndrome: the structure of subjective complaints after mTBI. J Head Trauma Rehabil 1995; 10: 1–17.

29 Weathers F , Litz B, Herman D, et al. : The PTSD Checklist (PCL): Reliability, Validity, and Diagnostic Utility. Annual Convention of the International Society for Traumatic Stress Studies; October 1993; San Antonio, TX.

30 Tombaugh TN : Test of Memory Malingering (TOMM) . New York, Multi-Health Systems, Inc., 1996.

31 Wilkinson GS , Robertson G: Wide Range Achievement Test (WRAT4) . Lutz, FL, Psychological Assessment Resources, 2006.

32 Dawson DR , Anderson ND, Burgess P, Cooper E, Krpan KM, Stuss DT: Further development of the Multiple Errands Test: standardized scoring, reliability, and ecological validity for the Baycrest version. Arch Phys Med Rehabil 2009; 90(11 Suppl): S41–51.

33 Stern RA , White T: Neuropsychological Assessment Battery . Lutz, FL, Psychological Assessment Resources, 2003.

34 Reynolds CR : Comprehensive Trail Making Test (CTMT) . Austin, TX, Pro-Ed, 2002.

35 Harris PA , Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG: Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009; 42(2): 377–81.

36 Jones A : Test of memory malingering: cutoff scores for psychometrically defined malingering groups in a military sample. Clin Neuropsychol 2013; 27(6): 1043–59.

37 Baum CM , Connor LT, Morrison T, Hahn M, Dromerick AW, Edwards DF: Reliability, validity, and clinical utility of the executive function performance test: a measure of executive function in a sample of people with stroke. Am J Occup Ther 2008; 62(4): 446–55.

38 Cohen J : Statistical Power for the Behavioral Sciences , Ed 2, Hillsdale, NJ, Erlbaum Associates, 1988.

39 Morrison MT , Giles GM, Ryan JD, et al. : Multiple Errands Test–Revised (MET–R): a performance-based measure of executive function in people with mild cerebrovascular accident. Am J Occup Ther 2013; 67(4): 460–8.

40 Wolf TJ , Morrison T, Matheson L: Initial development of a work-related assessment of dysexecutive syndrome: the Complex Task Performance Assessment. Work 2008; 31(2): 221–8.

41 Wolf TJ , Barbee AR, White D: Executive dysfunction immediately after mild stroke. OTJR (Thorofare NJ) 2011; 31(1): S23–29.

42 Montgomery GK : A multi-factor account of disability after brain injury: implications for neuropsychological counselling. Brain Inj 1995; 9(5): 453–69.

43 Schutz LE , Wanlass RL: Interdisciplinary assessment strategies for capturing the elusive executive. Am J Phys Med Rehabil 2009; 88(5): 419–22.

44 Boyle E , Cancelliere C, Hartvigsen J, et al. : Systematic review of prognosis after mild traumatic brain injury in the military: results of the International Collaboration on Mild Traumatic Brain Injury Prognosis. Arch Phys Med Rehabil 2014; 95(3 Suppl 2): S230–7.

45 Aupperle RL , Melrose AJ, Stein MB, et al. : Executive function and PTSD: disengaging from trauma. Neuropharmacology 2012; 62(2): 686–94.