AWARD NUMBER: W81XWH-16-1-0581
BA150325

TITLE: Development, Reliability, and Equivalence of an Alternate Form for the CQ Duty Performance-based Measure

PRINCIPAL INVESTIGATOR: Mary Vining Radomski

CONTRACTING ORGANIZATION: Allina Health

REPORT DATE: December 2020

TYPE OF REPORT: FINAL REPORT

PREPARED FOR: U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	**2. REPORT TYPE**	**3. DATES COVERED**
December 2020	Final Report	30 Sep 2016- 29 Sep 2020

4. TITLE AND SUBTITLE

Development, Reliability, and Equivalence of an Alternate Form for the CQ Duty Performance-based Measure

5a. CONTRACT NUMBER
W81XWH-16-1-0581

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

Mary Vining Radomski

E-Mail: mary.radomski@allina.com

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Allina Health
800 E. 28th Street
Minneapolis, MN 55407-3723

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

U.S. Army Medical Research and Development Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Previous research demonstrated that the recently-developed Charge of Quarters Duty Test (CQDT), a performance-based assessment of executive function, can be reliably administered and distinguishes between known-groups of healthy control soldiers and those with traumatic brain injury. As such, the CQDT shows promise in helping to inform readiness to return to duty and need for rehabilitation. However, performance based assessments that involve multitasking such as the CQDT, cannot be repeated as a post-treatment outcome measure due to learning effects. Therefore, an equivalent alternate form is needed. To address this gap, a team of rehabilitation researchers and clinicians used a systematic process to develop a presumed alternate form; evaluated its equivalence relative to the CQDT; examined a larger data set in order to propose preliminary score interpretation. They achieved the following results. Using systematic, iterative process the study team developed the Maintenance Office Duty Test (MODT), the alternate form for the CQDT. Sixty-two participants enrolled in the study, 57 of whom completed all study procedures. 40 participants performed both versions of the tests as well as a series of neurocognitive measures. Examination of participants' test scores and experiential ratings on both versions suggests that the MODT and CQDT are equivalent. Exploratory analyses of a dataset of 210 cases (CQDT, MODT, FDDT [a civilian version of the CQDT) suggested that test scores in the lowest decile may signify impaired performance. Further study is needed.

15. SUBJECT TERMS

Mild traumatic brain injury, performance-based assessment, multitasking, executive functions, alternate form

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18. NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT	**b. ABSTRACT**	**c. THIS PAGE**			USAMRMC
Unclassified	Unclassified	Unclassified	Unclassified	45	**19b. TELEPHONE NUMBER** *(include area code)*

TABLE OF CONTENTS

Page

INTRODUCTION

Many Service Members (SM) experience concussion (also known as mild traumatic brain injury [mTBI]) as a result of military combat and training, motor vehicle crashes, and sports-recreational activities. After mTBI, SM may experience an array of sensorimotor and cognitive problems, including difficulty with executive functions. Executive functions refer to higher order thinking abilities that enable goal-directed behavior, particularly in novel situations where people lack well-learned behaviors to draw upon. Because there are evidence-based cognitive rehabilitation interventions that can improve executive functioning, it is important to identify SM with mTBI with executive dysfunction who should be referred to these services before resuming high-consequence activities such as military duty. Unfortunately, executive dysfunction often goes undetected because traditional neurocognitive measures are designed to evaluate single domains rather than integrated functioning and the high levels of structure inherent in these assessments fail to adequately challenge the impaired functions. Performance-based assessment requires the patient to perform tasks that simulate the demands of everyday activities while the examiner uses behaviorally-based metrics to quantify functioning. **Performance-based measures designed to incorporate multitasking appear to be particularly sensitive to detecting deficient executive functions**. Existing performance-based involving multitasking have demonstrated sensitivity to executive dysfunction but the nature of the task components may lack face validity for SM with mTBI and their superiors, especially as related to readiness for return to duty.

Previous research demonstrated that the recently-developed Charge of Quarters Duty Test (CQDT), a performance-based assessment of executive function, can be reliably administered and distinguishes between known-groups of healthy control soldiers and those with traumatic brain injury. As such, the CQDT shows promise in helping to inform readiness to return to duty and need for rehabilitation. However, performance based assessments that involve multitasking such as the CQDT, cannot be repeated as a post-treatment outcome measure due to learning effects. Therefore, an equivalent alternate form is needed.

Research Question #1. Can an expert team of military, Veterans Administration, and civilian rehabilitation researchers and clinicians develop an equivalent alternate form of the CQD that is experienced as novel by SM with mTBI and that can be scored reliably by 2 independent raters?

Technical Objective #1: Develop an alternate form of the CQD.

Research Question #2. To what extent is the CQD-AF equivalent to the CQD-O based on a) difference of paired scores for both forms of the CQD and b) correlation between participants' performance of neurocognitive measures of executive functioning and each version of the CQD?

Technical Objective #2: Evaluate equivalence of CQD-AF.

Research Question #3: To what extent can a group of researchers and expert clinicians use empirical, theoretical, and clinical evidence to come to consensus regarding multitasking test score patterns that signify possible executive dysfunction?

Technical Objective #3: Propose scoring interpretation guidance

KEYWORDS

Mild traumatic brain injury
Executive function
Performance-based assessment
Multitasking
Alternate form

ACCOMPLISHMENTS

Note: Study findings are detailed in Appendix A and summarized in the table below.

Major goals of the project	Accomplishments under these goals
Use an expert team to develop an alternate form (AF) for the CQDT that can be scored reliably	• Assembled an expert team comprised of rehabilitation researchers and clinicians who have experience with multitasking test development and/or clinical testing and intervention for executive dysfunction associated with mTBI • Conducted detailed task analyses to specify demands of the CQDT; this involved observing people perform the test; administering and taking the test • Used a systematic process to develop 4 possible alternate forms, which were then evaluated by the expert team; the 2 options deemed to have the most potential were reduced to practice • Developed 2 options for the AF, which were then evaluated (task analyses, experiential evaluation by administering and taking the 2 options) • Used a consensus process to select the Maintenance Office Duty Test (MODT) as the CQDT's alternate form (see details in Appendix B) • Verified sufficient scoring agreement for the MODT-CQDT for 2 raters through an iterative process. Scoring by both raters established inter-rater reliability based on pre-established tolerances (Total Performance Times are within ± 1 minute of each other; Total Visits scores are within ± 1 of each other; Total Task Performance Accuracy scores are within ± 2 of each other). Ten participants (3 with a history of acquired brain injury) performed either the CQDT or MODT twice, approximately 2 weeks apart. This resulted in 60 observations (20 test administrations x 3 test metrics); 58/60 observations were within a priori tolerances for rater agreement.
Evaluate equivalence of CQDT-AF	• Sixty-two participants enrolled in the study, 57 of whom completed all study procedures. Data from 40 participants were used to evaluate the equivalence of the CQDT and MODT, 16 people with a history of ABI and 24 who did not (Healthy Controls [HC]). • To evaluate equivalence, participants performed the 2 test versions approximately 2 weeks apart. We compared mean scores and standard deviations for both forms of the test based on the hypothesis that if the tests are equivalent, subscores should not be significantly different. We found that there were no statistically significant differences in subscores on the MODT and CQDT, suggesting that the 2 tests were likely equivalent. • We also evaluated equivalence by examining whether or not MODT and CQDT subscores were correlated with neurocognitive test subscores, which we hypothesized to challenge similar cognitive domains as the multitasking tests. We hypothesized that if the 2 tests are equivalent, the correlations between various MODT subscores and various neurocognitive measures should be similar to that of the CQDT and various neurocognitive measures. However, we also presumed that these correlations would be weak at best, consistent with similar evaluations of other performance-based tests of multitasking. Overall, correlations between MODT and CQDT and neurocognitive test subscores were weak, with a few exceptions where correlations were statistically significant and moderately weak to moderate. It is notable that both the MODT and CQDT had statistically significant correlations between Performance Accuracy and two MIST subscores, which measure prospective memory, two NIH Toolbox tests (Cognition Fluid Composite and Cognition Total Composite Score). • Finally, we explored the equivalence of the 2 tests by examining how participants rated their experience with test demands. Immediately upcoming completing one of the multitasking tests, participants filled out a modified NASA experience questionnaire in which they rated the extent to which they experienced various task domains as challenging. We hypothesized that if the 2 tests are equivalent, participants' ratings related to test demands should be similar for both tests. There were no statistically significant differences in challenge ratings for the 2 tests
Use an expert consensus process to propose	• We assembled an aggregate dataset from 3 studies in which 3 versions of the same multitasking test were administered to healthy controls and individuals with ABI (CQDT, MODT, and Front Desk Duty Test [FDDT], a civilian version of the CQDT). We then

scoring interpretation guidance for the CQDT and CQDT-AF	conducted an array of exploratory, iterative analyses in order to propose preliminary cut-points that might suggest impaired test performance. • After examining a series of score distributions, we conducted analyses using the Total Visits score to evaluate the utility of using scores at the 90th percentile to identify problematic multitasking performance. We selected this particular subscore as the "master predictor" because it represents the primary hypothesized challenge of the test, that being to figure out how to complete the test assignments in as few visits as possible. Based on scores from 210 participants: mean number of visits = 16.40; median number of visits = 14.00; 90th percentile value = 27. • We further evaluated mean subtest scores based on whether or not participants' # of visit fell at or below the 90th percentile (1 group in Table 17). Mean subtest scores were compared based on 2 groups (0=# of visits > 90th percentile; 1 = # of visits ≤ 90th percentile) using the Independent Samples Test; there were statistically significant group differences (p=.000 for each - total visits, rule break count, total time, prospective memory score). • These findings suggested that when a person performs the multitasking test and has a Total Visits score in the 90th percentile, the other sub-scores also likely shift in a negative direction as well. Further, the other subtest scores are significantly different when comparing scores of people whose Total Visits fell in the 90th percentile versus those whose Total Visits fell in the more normal range. • These analyses affirm the potential utility of using test score values in the 90th percentile to identify problematic test performance. The research team concluded that if a participant has 2 subscores at the 90th percentile, clinicians should consider further treatment or evaluation for executive dysfunction. However, performance of complex multitasking tests such as the CQDT, MODT, and FDDT should never be evaluated solely by subtest scores. Clinicians are advised to consider behavioral observations during test performance such as the patient's non-verbal responses starting at the time of presenting the instructions.

What opportunities for training and professional development has the project provided?

Nothing to Report as yet. We had planned to provide a training workshop at Courage Kenny Rehabilitation Institute in April 2020 but this was cancelled due to COVID-19.

How were the results disseminated to communities of interest?

Interim progress was reported in 4 presentations:

- American Occupational Therapy Association Annual Conference – April 2018
- Interim Progress Report (USAMRMC) – April 2018
- Military Health System Research Symposium – August 2019
- Minnesota Occupational Therapy Association Annual Conference - October 2019

What do you plan to do during the next reporting period to accomplish the goals?

This is the Final report – Nothing to Report.

However, in 2021 (after study closure), the research team will continue to implement their dissemination plan. This includes developing and submitting a manuscript for publication and making test materials available to those interested via the Courage Kenny Rehabilitation Institute website. Military and civilian stakeholders at DVBIC and Veterans Administration will be notified of the availability of these materials in order to advance development of follow-on research and utilization.

IMPACT

What was the impact on the development of the principal discipline(s) of the project?

Nothing to Report as yet. However, the methodology used to successfully develop and evaluate an equivalent alternate form for a multitasking test might be replicable for other research groups in the future.

What was the impact on other disciplines?

Nothing to Report.

What was the impact on technology transfer?

Nothing to Report.

What was the impact on society beyond science and technology?

Nothing to Report.

CHANGES/PROBLEMS

Changes in approach and reasons for change

Nothing to Report.

Actual or anticipated problems or delays and actions of plans to resolve them

In early 2020, we had to discontinue subject enrollment and data collection due to COVID-19. However, this had no negligible effects on study findings.

Changes that had a significant impact on expenditures

Nothing to Report.

Significant changes in use or care of human subjects, vertebrate animals, biohazards, and/or select agents

Nothing to Report.

PRODUCTS

Publications, conferences, and presentations

Interim progress was reported in 4 presentations at 3 conferences and 1 IPR.

- American Occupational Therapy Association Annual Conference – April 2018
- Interim Progress Report (USAMRMC) – April 2018

- Military Health System Research Symposium – August 2019
- Minnesota Occupational Therapy Association Annual Conference - October 2019

A manuscript will be developed and submitted for publication in 2021.

Website or Internet site

Nothing to Report.

Technologies or techniques

We developed an equivalent alternate form for the CQDT called the Maintenance Office Duty Test (MODT). The test is detailed in Appendix B.

Inventions, patent applications and/or licenses

Nothing to Report.

Other products

Nothing to Report.

PARTICIPANTS & OTHER COLLABORATING ORGANIZATIONS

What individuals worked on the project?

Name	Mary Vining Radomski
Project Role	PI/PD
Researcher Identifier	0000-0003-0600-4494
Nearest person month worked	1 month
Contribution to the Project	Overseeing all aspects of the study; developing IRB submissions; developing progress reports; interpreting findings
Funding Support	This grant

Name	Ginger Carroll
Project Role	Research Coordinator
Researcher Identifier	0000-0001-9031-991X
Nearest person month worked	1 month
Contribution to the Project	Recruiting, screening, consenting participants; overseeing data collection and data management
Funding Support	This grant

There were no other individuals who contributed at least one person month over the past year. However, many researchers and clinicians contributed to the study including the following individuals.

Courage Kenny Research Center
Rob Kreiger, PhD
Kristina Kath, BS

Courage Kenny Rehabilitation Institute
Amy Meyers, OTR/L
Chris Tripp, MS, OTR/L

Colorado State University
Aaron Eakman, PhD, OTR/L

Leslie Davidson, PhD, OTR/L

Tim Wolf, PhD, OTR/L

Year 1 -2 only: Laurel Smith, MBA, OT/L; Joette Zola, OTR/L

Has there been a change in the active other support of the PD/PI or senior/key personnel since the last reporting period?

Nothing to Report.

What other organizations were involved as partners?

Nothing to Report.

SPECIAL REPORTING REQUIREMENTS

Collaborative Awards

Not applicable.

Quad Chart

See attached and also submitted separately to eBRAP.

Development, Reliability, and Equivalence of an Alternate Form of the CQ Duty Performance-based Measure

ERMS/Log Number: BA150325

W81XWH-16-1-0581

PI: Mary Vining Radomski, PhD, OTR **Org:** Allina Health/Courage Kenny Research Center **Award Amount:** $247,961.00

Study Aims

Technical Objective #1: Develop an alternate form of the CQD.
Technical Objective #2: Assure rater agreement across 2 raters in administering the CQDT-Original (CQDT-O) and CQDT-Alternate Form
Technical Objective #3: Evaluate equivalence of CQDT-AF based on a) difference of paired scores for both forms of the CQDT and b) correlation between participants' performance of neurocognitive measures of executive functioning and each version of the CQDT .

Approach

This study involves 2 parts. In part 1, a team of rehabilitation research experts from military, VA, civilian sectors will conduct task analyses of the CQDT-O in order to characterize key demands. Based on the results, the team will develop 2 workable CQDT-AF, conduct similar task analyses, and select the option that appears to be equivalent to the CQDT-O. In part 2, up to 60 individuals (some with a history of mild TBI) will be recruited. Inter-rater agreement will be evaluated by 2 raters on the first 10-15 subjects. All subjects will perform CQDT-O, CQDT-AF, and neurocognitive measures to assess equivalence and construct validity.

Performance-based measures that involve multitasking appear to be sensitive to deficient executive functions that can occur with TBI; this type of test can't be repeated due to practice effects. The CQD Test (CQDT), part of the recently developed *Assessment of Military Multitasking Performance*, was found to be reliable and differentiate between healthy controls and SM with mild TBI. **If we succeed in developing an equivalent alternate form, the CQDT may be used to both identify executive dysfunction and quantify treatment outcomes in SM with mTBI**.

Accomplishments (Yr 4, Q16): Data analyses were completed. Overall finding: The alternate form for the CQDT (the Maintenance Office Duty Test [MODT]) can be scored reliably and appears to be equivalent to the CQDT.

Timeline and Cost

Activities CY	16	17	18	19	20
Establish contracts; prepare for task analyses; submit Part 1 protocol to Allina IRB					
Develop CQDT-AF					
Conduct validation study on CQDT-AF					
Data analysis and dissemination					
Estimated Budget ($K)	**$10**	**$40**	**$90**	**$70**	**$50**

Updated: 12/22/2020

Goals/Milestones

CY16 Goal – Study kick off
☑ Functionality tests of integrated firmware and software
☑ Development and submission of Part 1 protocol to Allina IRB
☑ Obtain approval from Part 1 protocol to USAMRMC HRPO

CY17 Goals – Develop CQD-AF
☑Conduct task analyses to characterize CQD-O
☑Develop CQD-AF options

CY18 -19 Goal –Collect CQDT-CQDAF data
☑Select best CQDT-AF option to subject to validation
☑Part 2 protocol approved by AH 6-7-18 and by HRPO 10-9-18
☑Subject enrollment, data collection for validation study (part 2)
☑Complete subject enrollment, data collection for validation study

CY2020 Goals-Complete validation study; analyze/report findings
☑Finalize data collection
☑Finalize data analysis

Budget Expenditure to Date
Projected Total Expenditure: $247,961.00
Actual Expenditure to date: $247,961.00

APPENDICES

APPENDIX A: STUDY FINDINGS

Development, reliability, and equivalence of an Alternate Form of the CQ Duty Performance-based Measure (#W81XWH-16-1-0581)

The purpose of this award was to develop and evaluate the equivalence of Charge of Quarters Duty Alternate Form (CQDT-AF) by answering the following research questions.

- Can an expert team of military, Veterans Administration, and civilian rehabilitation researchers and clinicians develop an equivalent alternate form of the CQD that is experienced as novel by SM with mTBI and that can be scored reliably by 2 independent raters?
- To what extent is the CQD-AF equivalent to the CQD-O based on a) difference of paired scores for both forms of the CQD and b) correlation between participants' performance of neurocognitive measures of executive functioning and each version of the CQD?
- To what extent can a group of researchers and expert clinicians use empirical, theoretical, and clinical evidence to come to consensus regarding multitasking test score patterns that signify possible executive dysfunction?

In this report, we report the study findings in 2 parts. In Part 1, we summarize the development of the CQDT-AF and evaluation of scoring reliability. In Part 2, we summarize our evaluation of the equivalence of the CQDT and the CQDT-AF and propose scoring interpretation guidance.

PART 1: DEVELOPMENT OF THE CQDT-AF

Research Question: Can an expert team of military, Veterans Administration, and civilian rehabilitation researchers and clinicians develop an equivalent alternate form of the CQD that is experienced as novel by SM with mTBI and that can be scored reliably by 2 independent raters?

INTRODUCTION

In Part 1 of the funded study, a team of rehabilitation researchers and clinicians (one military clinician) used an iterative process in the hopes of developing a CQDT-AF that had similar test challenges as the CQDT and yet would be experienced as novel by participants. This process had 3 phases: characterizing the demands of the CQDT; specifying an "optimal" CQDT-AF; measuring scorer agreement.

1.0 Characterize the demands of the CQDT

In Year 1, we developed a task analysis framework in an effort to characterize the task load of the CQDT, the original version of the test. The plan was to obtain team consensus on the CQDT task load in order to replicate the task load in designing an equivalent CQDT-AF. To that end, we first conducted experiential analyses in which each of the 6 occupational therapy investigators were administered the test. They then completed the task analysis form to describe their experience of task load. Experiential data from the 6 occupational therapists were analyzed by examining interrater reliability using Krippendorff's Alpha, which suggested that the 6 raters had different CQD task-load score patterns. Further analyses using box plots based on raters, task components, and task constructs continued to suggest task load largely varied by individual rather than task element. Overall, raters had some agreement that task load was greatest for cognitive elements and lowest for environmental demands. These analyses resulted in refinement of the task analysis framework/form so as to enhance clarity of task component definitions and rating scores.

Next we conducted observation-based task analyses of the CQD-O. CKRC investigators administered the CQD-O to two adults (one who was healthy, 1 with a history of ABI). Their performance was videorecorded and the 6 occupational therapy investigators reviewed the videorecorded performances and again used the task analysis to characterize the CQD-O task load. Efforts to develop a granular, singular task demand characterization using these methods were impeded by obvious participant differences in what aspects of the task they prioritized. For example, the healthy individual appeared to focus prioritize making as few trips to work stations to complete the task, ask stated explicitly in the instructions. Thus, this participant appeared to experience significant task press associated with forethought and planning. However, the individual with a history of ABI appeared to focus on memory elements of the task, seemingly neglecting the instruction to make as few visits to work stations as possible (the planning component). These observations called into question the ability to characterize the task demands of the CQDT separate from how individuals uniquely experience the task, causing us to abandon our goal of developing a single ubiquitous, consensus-based task characterization of the CQD-O as preliminary to developing options for the CQD-AF. Instead, Tim Wolf, PhD, OTR/L developed a rubric that

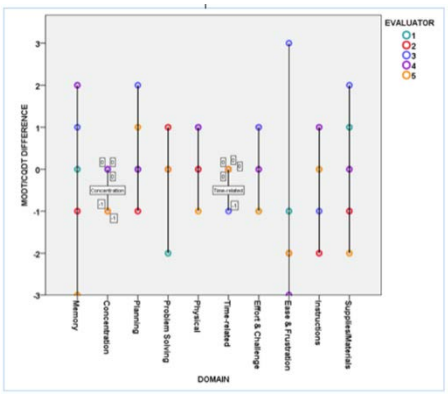
operationalized various elements of the CQD-O, which was used as the foundation for developing options for the CQD-AF.

2.0 Specify an optimal CQD-AF

In Year 1, the research team used a systematic process to specify the cognitive and sensorimotor demands of the CQDT (the original version of the test). In Year 2, these findings were used to inform a test development process in which 2 sub-teams were each charged with designing at least 2 alternate form options, 4 total (multitasking contexts: fuel-check maintenance office; briefing scenario; camping; standing up a FOB market). Each of these options were subjected to task analysis using the same framework used in analyzing the CQDT and the Burgess definition of multitasking. During an all-team TCON on December 4, 2017, all 4 options were presented and evaluated; the team determined that Fuel-check Maintenance and Camping CQDT-AF options offered the most potential as logistically feasible, reasonable face validity, and potential equivalence with the CQDT. Those 2 test scenarios were reduced to practice and 2 participants' performances of both test options were videorecorded and analyzed by the research team.

On January 8, 2018, the entire team participated in an all-day work session to evaluate the 2 best CQD-AF options and select the one with the most promise for equivalence. Team members used experiential task analyses, observation, and discussion to come to consensus on the alternate form to advance to equivalence testing – the Maintenance Office Duty Test (MODT). Subsequently, 5 occupational therapists performed both CQDT and MODT, rating challenge associate with various tasks domains. Results were used to finalize MODT refinements; experiential analyses suggesting that CQDT and MODT have similar task challenges (see Figures 1 & 2).

Figure 1. Demand impressions



Notes. Close agreement on the Concentration and Time-Related dimensions with respect to comparative demands for the two test versions; wide variation among rater impressions for Ease & Frustration, Memory, and Supplies/Materials dimensions; intermediate levels of rater agreement for Planning, Problem Solving, Instructions

Figure 2. Krippendorff's alpha rater agreement results for CQDT and MODT suggest moderate rater agreement of comparative demands.

	Alpha	LL95%CI	UL95%CI	Units	Observers	Samples
CQDT	.4486	.2964	.5867	10	5	10,000
MODT	.4607	.3279	.5849	10	5	10,000

3.0 Assure rater agreement between 2 raters

An iterative process was used to achieve scoring agreement between the 2 raters who collected data in Part 2 of this study. Scoring by both raters established inter-rater reliability based on pre-established tolerances (Total Performance Times are within ± 1 minute of each other; Total Visits scores are within ± 1 of each other; Total Task Performance Accuracy scores are within ± 2 of each other).

Ten participants (3 with a history of acquired brain injury) performed either the CQDT or MODT twice, approximately 2 weeks apart. This resulted in 60 observations (20 test administrations x 3 test metrics); 58/60 observations were within a priori tolerances for rater agreement.

PART 2: EVALUATING THE EQUIVALENCE OF THE MODT AND CQDT AND PROPOSING INTERPRETATION GUIDANCE

INTRODUCTION

In Part 2 of the funded study, investigators answered two fundamental research questions: 1) Is the Maintenance Office Duty Test (MODT) an equivalent form of the Charge of Quarters Duty Test? and 2) To what extent can a group of researchers and expert clinicians use empirical, theoretical, and clinical evidence to come to consensus regarding multitasking test score patterns that signify possible executive dysfunction? Each primary question is informed by a series of sub- and exploratory questions, which together inform study conclusions.

1.0 Research Question: Is the Maintenance Office Duty Test (MODT) an equivalent alternate form of the Charge of Quarters Duty Test (CQDT)?

1.1 Methods

1.1.1 Study Design

Data were collected to both evaluate inter-rater reliability of the MODT and to test the equivalence of the CQDT and MODT (Figure 1). The equivalence portion of the study employed a repeated measures design in which each participant was tested under all conditions and acts as his or her own control. The following analyses were conducted in order to inform answers to the research question:

- Equivalence of the MODT-CQDT based on the examination of paired differences in performance of the task metrics (task performance accuracy, total performance time, total number of transits) for the CQDT (section 1.2.3)
- Equivalence of the MODT-CQDT based on correlations of both versions with neurocognitive measures (section 1.2.4) - see Table 1 for list of measures other than CQDT or MODT.
- Equivalence of the MODT and CQDT based one whether or not participants who performed both tests experienced them similarly (section 1.2.5).
- (Exploratory) Extent to which learning effects are minimized when participants take 2 different test versions (CQDT-MODT) versus taking the same test twice (1.2.6).
- (Exploratory) Extent to which performance on multitasking tests (CQDT or MODT) was explained by dynamic personal factors (e.g. self-reported fatigue, pain, distraction) (section 1.2.7).

Figure 1. Overview of study procedures

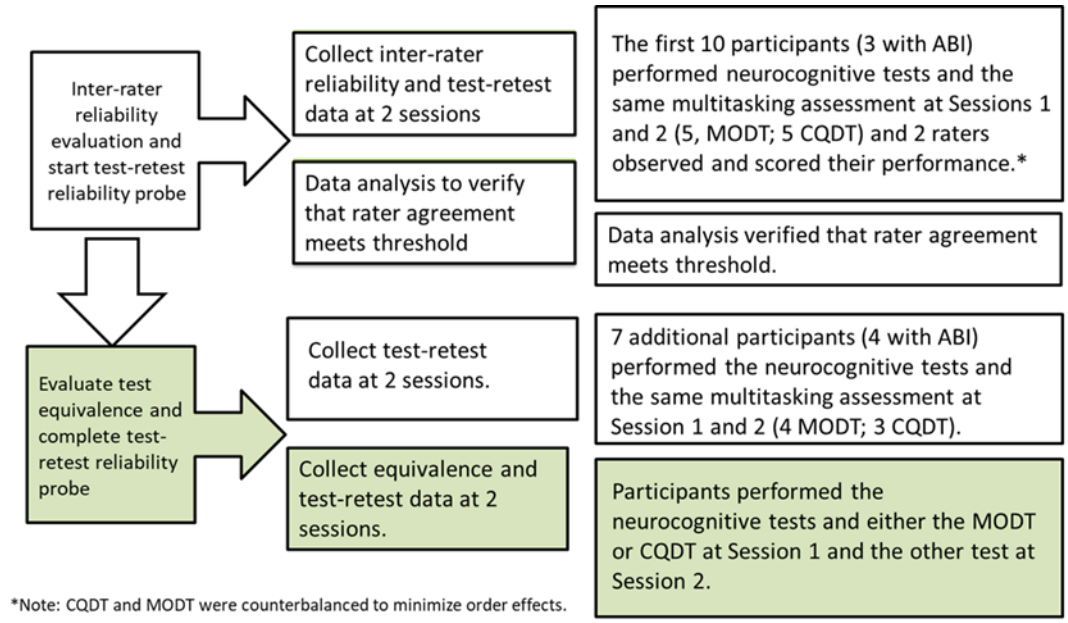


Table 1. Measures/Questionnaires

Instrument	Description
How are you doing today? Questionnaire	A study-developed self-report questionnaire in which participants describe general pain, fatigue status on day of testing.
Experience Questionnaire – Revised	A self-report questionnaire modified from NASA in which participants describe the extent to which they experienced various task domains as challenging.

Wide Range Achievement Test (WRAT 4) Reading Test	A standardized reading test that will be used to establish an estimate of intelligence. The reading test consists of 55 words that increase in complexity and atypical pronunciation.
Cognitive battery from the NIH Toolbox	A 25-30 minute test of the mental processes involved in gaining knowledge and comprehension, such as thinking, knowing, remembering, judging, and problem-solving. These tests will be downloaded onto an iPad using the NIH Toolbox App; participants perform the tests on an iPad.
Tower of Hanoi (TOH)	A test that challenges foresight and planning ability by requiring participants to arrange small discs to conform to pictorial examples while adhering to task rules. We are using the TOH from the Delis-Kaplan Executive Function System (DKFES)
Memory for Intentions Screening Test (MIST)	A performance-based measure of prospective memory, which allows for comparisons of performance with event-based and time-based cues involving short (2 minute) and long (15 minute) delays.

1.1.2 Study Activities

Participants attended 2 sessions at either the Courage Kenny Research Center (CKRC) or Mercy Hospital. At the first session, each participant performed a set of neurocognitive tests and one of the multitasking assessments, either the CQDT or the MODT. Within 1 to 3 weeks, the participant returned to the same testing location and performed the other set of neurocognitive tests and either the CQDT or MODT (Box 1).

Box 1. Overview of study sessions

- At each session, participants completed a status questionnaire ("How Are You Doing Today?") followed by the neurocognitive measures.
    - At Session 1, participants performed the WRAT-4 Reading test followed by the MIST.
    - At Session 2, participants performed the NIH Cognitive battery followed by the TOH.
- After completing the neurocognitive measures, participants were offered a 5-10 minute break.
- The participant then performed one of the multitasking assessments (CQDT or MODT).
- After performing the multitask assessment, the participant filled out the Experience Questionnaire-Revised.
- The participant received a reminder about the date/time of Session 2 within a few days of the session.
- Session 2 followed the same sequence as described for Session 1.
- Upon completing Session 2, the participant was issued a $50 gift card.

1.1.2 Participants

The study was approved to recruit a convenience sample of up to 80 adult civilians, Reservists, National Guard members, and military veterans, with and without a history of acquired brain injury (ABI) from Allina Health and the surrounding Minneapolis-St. Paul community.

Inclusion criteria for participants with self-reported history of ABI:

- At least 6 months post a self-reported history of ABI such as from stroke, concussion, traumatic brain injury, anoxia, hemorrhage
- 18 – 55 years of age
- Able to read and speak in English
- Graduated from high school (or equivalent)
- Able to communicate (express themselves and comprehend) abstract information (e.g., related to discharge planning, current events, religion, humor, math, finances used in daily life)
- Ambulatory with or without assistive device
- Able to hear with or without assistive device
- Able to use a pen or pencil to write information on a piece of paper

Exclusion criteria:

- Lacks bilateral hand function
- Self-reported history of neurological disease

- Self-reported unmanaged mental illness
- Self-reported drug/alcohol abuse
- Inability to write, comprehend verbal instructions
- Profound vision loss such that the subject cannot read instructions written in 12 point font (with or without correction) or identify objects in the testing space
- Self-reported dementia
- Unable to arrange transportation to the testing session

Inclusion criteria for participants without a history of acquired brain injury:

- Ages 18 – 55 years of age
- Able to read and speak in English
- Graduated from high school (or equivalent)
- Able to communicate (express themselves and comprehend) abstract information (e.g., related to discharge planning, current events, religion, humor, math, finances used in daily life)
- Ambulatory with or without assistive device
- Able to hear with or without assistive device
- Able to use a pen or pencil to write information on a piece of paper

Exclusion criteria:

- Self-reported history of ABI from stroke, traumatic brain injury, anoxia, hemorrhage
- Lacks bilateral hand function
- Self-reported history of neurological disease
- Self-reported unmanaged mental illness
- Self-reported drug/alcohol abuse
- Inability to write, comprehend verbal instructions
- Profound vision loss such that the subject cannot read instructions written in 12 point font (with or without correction) or identify objects in the testing space
- Self-reported dementia
- Unable to arrange transportation to the testing session

1.2 Results

1.2.2 Participants

Sixty-two participants enrolled in the study, 57 of whom completed all study procedures. Data from 40 participants were used to evaluate the equivalence of the CQDT and MODT, 16 people with a history of ABI and 24 who did not (Healthy Controls [HC]). See Table 2.

Table 2. Numbers of participants

Participant category	Total n	Completed protocol	Test-retest		Test equivalence 16 w ABI 24 HC	
			CQDT	MODT	CQDT > MODT	MODT> CQDT
Enrolled	**62**	**57**				
ABI	27	23	5	2	8	8
HC	35	34	3	7	10	14

We examined characteristics of study participants and any differences between participants with and without ABI (Table 3). In general, there were more female than male participants. Four of the 40 participants had a history of military service.

Race categories as depicted on our form were ambiguous. Only one respondent indicated Hispanic, but they also listed White. Only one respondent indicated “Other” as a category, but also indicated two additional race identities. Four respondents indicated multiple races. This response set does cannot produce p-values, since categories are not mutually exclusive. Thus, we removed the ambiguous categories by x-ing them out in the original table. The choice here is to: (1) just report counts without p-values (2) omit the “Other” and Hispanic categories if p-values are important – this is the

current table, or (3) redo this to include a “Multiple” category, report Hispanic as a 0, since it is for a “Multiple” case, convert the single Other case to “Multiple”, since that is more accurate.

It is interesting that there were ABI-HC group statistically significant differences in just 2 subtests of the 3 instruments administered (MIST summary score [p=.041] and NIH Toolbox Cognition Fluid Composite [p=.034]), suggesting that individuals with a history of ABI did not have severe residual cognitive deficits. Note that we did not analyze subgroup differences in BRIEF-A subscores because there were too few cases (administered only to the final nine subjects).

Table 3. Characteristics of participants (next 2 pages)

Characteristic	Group *M* (*SD*) or n (%)			Difference between groups
	All Participants (n=40)	Healthy Control (n=24)	Acquired Brain Injury (n=16)	
Age[a]	36.3 (8.9)	35.3 (8.9)	37.8 (8.9)	p=.398
Gender[b]				
Male	13 (32.5)	4 (16.7)	9 (56.3)	p=.015*
Female	27 (67.5)	20 (83.3)	7 (43.7)	
Race[b]				
Asian/Pacific Islander	2 (5.0)	2 (8.3)	0 (0)	p=.067
Black/African American	1 (2.5)	0 (0)	1 (6.3)	
Hispanic/Latino (ethnicity not exclusive)	x	x See above text	x	
Native American	2 (5.0)	0 (0)	2 (12.5)	
White/Caucasian	35 (87.5)	22 (91.7)	13 (81.2)	
Other (specify)	x	x	x	
Employment status[b]				
Full time	31 (77.5)	19 (79.1)	12 (75.0)	p=.251
Part time	5 (12.5)	2 (4.2)	3 (18.8)	
Retired	0 (0)	0 (0)	0 (0.0)	
Student	3 (7.5)	3 (12.5)	0 (0.0)	
Other	1 (2.5)	0 (0)	1 (6.2)	
Military service[b]				
Yes	4 (10.0)	2 (8.3)	2 (12.5)	p=1.00
No	36 (90.0)	22 (91.7)	14 (87.5)	
Baseline WRAT				
Raw score[c]	63.9 (3.0)	63.6 (3.4)	64.4 (2.4)	p=.596
Standard score[c]	105.5 (8.0)	104.9 (8.9)	106.4 (6.7)	p=.533
Percentile rank[c]	62.6 (17.4)	61.0 (19.3)	65.1 (14.3)	p=.533
Grade equivalent[b]	x	x	x	p=.140
MIST[c]				
Time cue raw score	7.1 (1.1)	7.3 (0.9)	6.8 (1.2)	p=.155
Event cue raw score	7.8 (0.7)	7.9 (0.4)	7.5 (0.9)	p=.138
Summary score	44.4 (3.9)	45.5 (3.1)	42.8 (4.5)	p=.041*
DKFES TOWER TEST[c]				
Total # of items	8.9 (0.4)	9.0 (0.0)	8.1 (0.5)	p=.154
Total # of moves	122.2 (31.4)	129.2 (32.7)	111.8 (26.9)	p=.083
Total achievement scaled score	11.2 (2.8)	11.3 (2.5)	11.1 (3.4)	p=.984
NIH TOOLBOX INSTRUMENTS[c]				
Flanker Inhibitory Control and Attention Test (Raw score)	8.4 (0.8)	8.6 (0.5)	8.2 (1.0)	p=.219
List Sorting Working Memory Test (Age corrected standard score)	98.4 (11.3)	98.5 (11.1)	98.1 (11.9)	p=.780
Dimensional Change Card Sort Test (Raw score)	8.7 (1.3)	8.9 (1.3)	8.4 (1.2)	p=.102
Cognition Fluid Composite – Ability to reason and think flexibly (Age corrected standard score)	105.8 (17.5)	110.1 (15.9)	99.3 (18.3)	p=.034*

Cognition Total Composite Score (Age corrected standard score)	111.0 (13.2)	113.8 (13.1)	106.7 (12.7)	p=.118

Notes. *p<0.05
[a] t-test for continuous data
[b] Fisher's exact test for categorical data
[c] Mann-Whitney U test for ordinal or continuous non-normal data

1.2.3 Equivalence of MODT and CQDT based on comparison of test subscores

Matched-pair t-tests were used to compare CQDT and MODT mean scores and standard deviations for various outcome sub-scores. Our hypothesis was that if the test formats were equivalent, sub-scores should not be significantly different. By using matched-pair analysis, we were able to control for any latent potential individual-level confounding factors that might impact method comparisons. As summarized in Table 4, there were no statistically significant differences in subscores on the MODT and CQDT.

Table 4. Comparison of means scores on both forms of the tests

TEST SUBSCORES	CQDT (n=40)		MODT (n=40)		p values for differences in mean sub-scores
	Mean (SD)	Range (min-max)	Mean (SD)	Range (min-max)	
# who completed test	n=40	n/a	n = 40	n/a	
TOTAL performance score (of 38)	33.2 (4.2)	21 (17-38)	33.9 (4.0)	17 (22-39)	.174
# of rules broken (of 4)	0.3 (0.5)	1 (0-1)	0.3 (0.5)	2 (0-2)	.767
TOTAL # of rules broken	0.9 (2.0)	10 (0-10)	2.3 (4.8)	24 (0-24)	.076
TOTAL # of visits	12.2 (7.6)	43 (5-48)	12.8 (5.5)	27 (6-33)	.453
Performance time (minutes, seconds)	20.2 (6.3)	25.9 (12.2-38.1)	21.2 (5.0)	24.1 (13.9-38.0)	.181

**p<.05*
CQDT = Charge of Quarters Duty Test; MODT = Maintenance Office Duty Test *M* = mean; *SD* = standard deviation

1.2.4. Equivalence based CQDT and MODT subscores correlations with subscores of measures of presumed convergent validity

MODT and CQDT subscores were correlated with neurocognitive test subscores that were hypothesized to be challenge similar cognitive domains as the multitasking tests. We hypothesized that if the 2 tests are equivalent, the correlations between various MODT subscores and various neurocognitive measures should be similar to that of the CQDT and various neurocognitive measures. However, we also presumed that these correlations would be weak at best, consistent with similar evaluations of other performance-based tests of multitasking. Non-parametric Spearman's rank correlations (rho) were calculated for this analysis. This method has the advantages of not requiring distribution normality or a linear relationship between the test variables.

Overall, correlations between MODT and CQDT and neurocognitive test subscores were weak, with a few exceptions where correlations were statistically significant and moderately weak to moderate. It is notable that both the MODT and CQDT had statistically significant correlations between Performance Accuracy and two MIST subscores, which measure prospective memory, two NIH Toolbox tests (Cognition Fluid Composite and Cognition Total Composite Score).

Table 5. Neurocognitive tests and domains

Neurocognitive Test	Domain	CQDT (n=40)				MODT (n=40)			
		Accuracy	# of rule breaks	Total visits	Performance time	Accuracy	# of rule breaks	Total visits	Performance time
MIST	Time cue raw score	.476**	-.281	-.168	-.342*	.385*	-.174	-.261	-.233
	Event cue raw score	-.017	.059	.129	-.003	.073	.089	-.105	-.228

	Summary score	.388*	-.192	-.077	-.267	.371*	-.090	-.262	-.285
DKFES Tower Test	Total # of items	.293	-.151	-.195	-.184	.010	-.092	-.141	-.103
	Total # of moves	.034	.139	.227	-.141	-.040	.103	-.154	-.274
	Total achievement scaled score	.303	-.182	-.307	-.173	.383*	-.218	-.101	-.036
NIH Toolbox	Flanker Inhibitory Control and Attention Test	-.073	-.017	-.061	-.111	.274	-.201	-.141	-.198
	List Sorting Working Memory Test	.234	-.054	-.100	-.013	.366*	-.342*	-.013	-.015
	Dimensional Change Card Sort Test	.302	.089	-.237	-.108	.354*	-.149	-.286	-.119
	Cognition Fluid Composite – Ability to reason and think flexibly	.354*	-.087	-.316*	-.151	.430**	-.301	-.198	-.038
	Cognition Total Composite Score	.354*	-.172	-.204	-.201	.413**	-.215	-.023	-.120

Note. *p<0.05; **p<0.01

Cells are highlighted where both multitasks had sub-scores were significantly correlated with a neurocognitive test.

1.2.5. Equivalence based on test demands as rated by participants

Immediately upcoming completing one of the multitasking tests, participants filled out a modified NASA experience questionnaire in which they rated the extent to which they experienced various task domains as challenging. We hypothesized that if the 2 tests are equivalent, participants' ratings related to test demands should be similar for both tests. There were no statistically significant differences in challenge ratings for the 2 tests (Table 6).

Matched-pair t-tests were again used to compare CQDT and MODT mean scores and standard deviations for various experience item scores. As in some previously described tests, our hypothesis was that if the test formats were equivalent, the experience item scores should not be significantly different between MODT and CQDT. By using matched-pair analysis, we were able to control for any latent potential individual-level confounding factors that might impact method comparisons.

Table 6. Comparison of mean scores related to experience of test demands for both forms of the test

Experience ratings related to test demands	CQDT (n=40)		MODT (n=40)		p-values for mean experience rating
	Mean (SD)	Range (min-max)	Mean (SD)	Range (min-max)	
Memory	6.8 (2.3)	8 (2-10)	6.6 (2.0)	7 (2-9)	.719
Concentration	6.8 (2.4)	8 (2-10)_	6.4 (2.5)	9 (1-10)	.466
Planning	8.0 (1.9)	7 (3-10)	8.0 (1.5)	6 (4-10)	.943
Problem Solving	7.1 (2.2)	8 (2-10)	7.1 (1.8)	7 (3-10)	.936
Time related	5.2 (2.4)	8 (1-9)	6.0 (2.3)	8 (2-10)	.113
Ease & Frustration	5.3 (2.4)	8 (1-9)	5.5 (2.2)	8 (1-9)	.673
Instructions	5.3 (2.7)	8 (2-10)	4.7 (2.7)	10 (0-10)	.358
Supplies & Materials	3.6 (2.4)	8 (1-9)	3.7 (2.5)	10 (0-10)	.745

Note. CQDT = Charge of Quarters Duty Test; MODT = Maintenance Office Duty Test *M* = mean; *SD* = standard deviation **p<.05*

1.2.6 Exploratory: Extent of learning effects when participants performed 2 test versions as compared to when they performed the same test twice

Forty participants performed the MODT and the CQDT, the test order was randomly assigned (Table 7). Twelve participants performed the same test twice, either the MODT twice or the CQDT twice (Table 8). Exact T-tests were used to compare Time 1 and Time 2 mean scores. Results suggest that differences in Time 1 and Time 2 were statistically significant only for performance accuracy, break counts, and performance time when the same test was performed twice.

Table 7. Time 1 and Time 2 scores on MODT and CQDT

Performance domain	Mean	N	SD	Dif	p
Performance accuracy Total	33.2 33.9	40 40	4.2 4.0	0.7	.174
Breaks Count	0.9 2.3	40 40	2.0 4.8	1.4	.076
Total Visits	12.2 12.8	40 40	7.6 5.5	0.6	.453
Performance Time (min)	20.2 21.2	40 40	6.3 5.0	1.0	.181

Table 8. Time 1 and Time 2 scores when the same test was performed twice

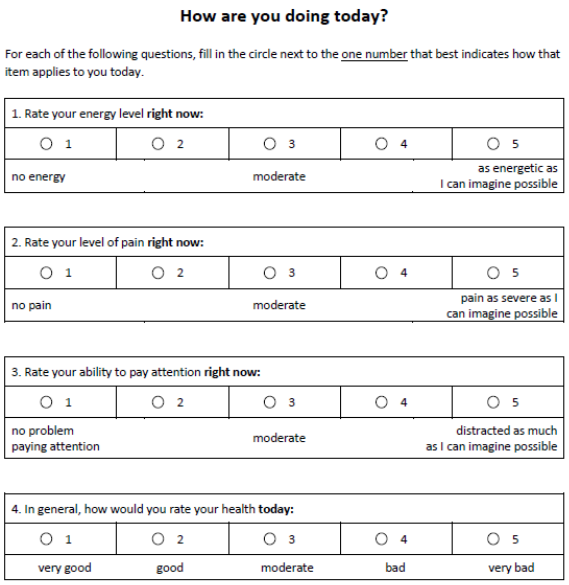
Performance domain	Mean	N	SD	Dif	p
Performance accuracy Total	33.6 35.8	12 12	3.9 2.2	2.2	.013*
Breaks Count	1.3 0	12 12	1.6 0	1.3	.015*
Total Visits	14.0 11.2	12 12	5.6 5.8	2.8	.110
Performance Time (min)	21.6 17.1	12 12	7.6 5.2	4.5	.002*

* $p<.05$

1.2.7 Exploratory: Extent to which multitasking performance is explained by dynamic personal factors (e.g., fatigue, pain, distraction)

During data collection, participants performed one of the 2 forms of the multitasking test. Prior to each test administration (n=116), participants completed a “How Are You Doing Today?” Questionnaire, comprised of 4 questions with response options in the form of Likert scales (see Figure 2).

Figure 2. How are you doing today questionnaire



How are you doing today?

For each of the following questions, fill in the circle next to the one number that best indicates how that item applies to you today.

1. Rate your energy level **right now:**				
○ 1	○ 2	○ 3	○ 4	○ 5
no energy		moderate		as energetic as I can imagine possible

2. Rate your level of pain **right now:**				
○ 1	○ 2	○ 3	○ 4	○ 5
no pain		moderate		pain as severe as I can imagine possible

3. Rate your ability to pay attention **right now:**				
○ 1	○ 2	○ 3	○ 4	○ 5
no problem paying attention		moderate		distracted as much as I can imagine possible

4. In general, how would you rate your health **today:**				
○ 1	○ 2	○ 3	○ 4	○ 5
very good	good	moderate	bad	very bad

Data distributions were non-normal for all dimensions of self-reported well-being. As summarized below, participants were generally not in pain, had adequate levels of energy, and viewed their health as acceptable, impressions which were stable across the 2 sessions. Overall, there was a somewhat greater score spread for "paying attention." See Tables 9 and 10.

Table 9. Summary of How Are You Doing Today? responses

Domains	**Session 1 (n= 59)**		**Session 2 (n=57)**	
	Mean	**SD**	**Mean**	**SD**
Energy level 1= No energy	3.46	.678	3.44	.824
Pain 1= No pain	1.37	.717	1.37	.762
Paying attention 1= no problem	2.17	1.053	2.33	.944
Overall health 1= very good	1.71	.617	1.81	.743

Table 10. Summary of self-reported concern levels in well-being domains at testing sessions

	Energy N (%)	**Pain**	**Paying Attention**	**Overall Health**
Not a concern*	108 (93.1%)	114 (98.3%)	64 (55.2%)	115 (99.1%)
Concern	8 (6.9%)	2 (1.7%)	52 (44.8%)	1 (.9%)

*On the 1-5 scale, 3 was the hinge value. Those who had concerns selected the 2 worse score options on the continuum.

Finally, we examined multitask test performance on Cases having two problematic responses on the How are You Doing survey, with score patterns described in Table 11. No real pattern here stands out for these cases. All of these participants indicated that Paying Attention was a concern, with

energy level the second problematic category in 6 of the 9 cases. Five of the 9 (55%) participants with concerns in 2 areas had a history of acquired brain injury.

Table 11. Multitasking test scores for participants who reported concerns in 2 domains of well-being at testing sessions

ID	Hx of ABI	Total_Score	Break_Count	Total_Visits	Time(Min)	Pros_Mem__Score
05CC	Y	31	2	22	17.7	5
06CC	Y	17	0	11	32.7	5
13MC	Y	37	0	8	22.5	8
21MC	N	37	0	10	19.8	7
39CM	N	35	0	7	18.3	7
40MC	Y	35	0	10	30	5
44MC	N	28	12	20	22.3	3
56MM	Y	38	1	11	20.4	8
57CC	N	30	0	18	19.4	3

1.3 Conclusions: The MODT is an equivalent alternate form of the CQDT. Using the 2 versions of this multitasking test (CQDT and MODT) appear to mitigate learning effects that may be introduced if participants perform the same version of the test twice. Results of exploratory analyses suggest that dynamic personal factors (such as self-reported energy level, pain, ability to pay attention, overall health) do not appear to influence multitasking test performance.

2.0 Research Question: To what extent can a group of researchers and expert clinicians use empirical, theoretical, and clinical evidence to come to consensus regarding multitasking test score patterns that signify possible executive dysfunction?

2.1 Methods

In this section, we summarize the exploration of empirical evidence to inform the effort to identify a multitasking test score performance pattern that signifies possible executive dysfunction.

2.1.1. Study Approach

We assembled an aggregate dataset from 3 studies in which 3 versions of the same multitasking test were administered to healthy controls and individuals with ABI (CQDT, MODT, and Front Desk Duty Test [FDDT], a civilian version of the CQDT). We then conducted an array of exploratory, iterative analyses. All 3 tests demonstrated sufficient inter-rater reliability in this or previous studies; also in previous studies the CQDT and FDDT demonstrated known-groups validity (Carroll et al., 2020; Radomski et al., 2018; Smith et al., 2014).

- First, we compared mean test subscores between HC and ABI participants.
- Next we examined score distributions, with a particular curiosity about values in the tale representing worse performance.
- Finally, we identified subscore values at the 90th percentile representing poor performance.

2.1.2 Participants

The aggregate dataset included records of 269 multitasking test scores performed by 212 individuals, 103 (~49%) of whom had a history of ABI.

Table 13. Number of test performances and participants by multitasking test in dataset

	Total	CQDT-AF study		FDDT study	AMMP study
		CQDT	MODT	FDDT	CQDT
Total test performances	269	57	57	70	85
Number of participants taking the tests	212	57		70	85

2.2 Preliminary exploration to identify score patterns suggestive of poor multitasking performance

2.2.1 Are there multitasking performance patterns signifying possible executive dysfunction based on differences in mean scores and extreme values between HC and participants with ABI?

2.2.1.1 Approach

In order to inform team discussions aimed at identifying sub-score patterns that were indicative of possible impaired performance, we first examined the differences in mean scores and "worst" values for healthy controls versus participants with ABI. Our hypothesis was that the extreme values of individuals with ABI would look "worse" in aggregate than the pattern for healthy control participants. Means/medians inform us if the distribution center values look different for ABI and non-ABI cases. Ranges tell us if the extreme value patterns look different for these groups.

2.2.1.2 Observations

Examining the medians and means in Table 14 shows there are only a few points of score difference between ABI and HC groups, regardless of the outcome being measured. This indicates no preponderant marginal differences between outcome averages relative to HC or ABI grouping. Evaluation of the extreme values shows some pattern difference for Total Accuracy, with ABI showing a lower value set for extremes. ABI also shows some higher extreme values for Total Time, with a very large single observation at 80.3 minutes. Further work would be needed to explore if singular extreme values represent individually-mediated factors in combination with ABI that produced the observed data pattern in extreme values.

Table 14. Multitask test score patterns by means, range, and extreme values

	ABI vs HC	**N**	**Mean**	**SD**	**Median**	**Range**	**5 most extreme values reflecting poor performance**
Total accuracy score (of 38)	HC	107	33.75	3.837	34.00	9-38	9, 24, 27, 28, 28
	ABI	103	31.73	5.371	33.00	10-38	10, 11, 17, 17, 20
Rule break count	HC	107	1.79	3.722	0	0-24	
	ABI	103	1.50	2.305	0	0-10	
Total visits	HC	107	15.37	8.072	13.00	6-53 visits	53, 44, 40, 39, 37
	ABI	103	17.47	10.512	14.00	6-53 visits	53, 52, 51, 51, 48
Total time	HC	107	19.49	5.211	18.3	11-43 min	40.2, 36.6, 30.8, 29.8, 29.1
	ABI	103	22.03	9.03	22.53	13 – 80 min	80.3, 47.2, 42.7, 39.7, 38.1
Pros mem score (of 8)	HC	107	6.60	1.446	7.0	2-8	2, 3, 3, 3, 3
	ABI	103	5.82	1.903	6.0	0-8	0,0,0,2,2

2.2.2 Are there multitasking performance patterns signifying possible executive dysfunction based on differences score distributions between HC and participants with ABI?

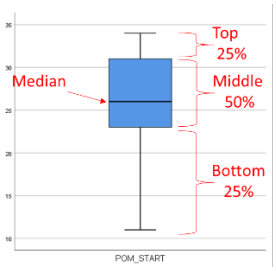
2.2.2.1 Approach

In order to inform team discussions aimed at identifying sub-scores that were indicative of possible impaired performance, we examined score distributions to determine if score patterns were observably different for healthy controls versus participants with ABI. Specifically, we inspected data histograms and box plots.

2.2.2.2 Observations

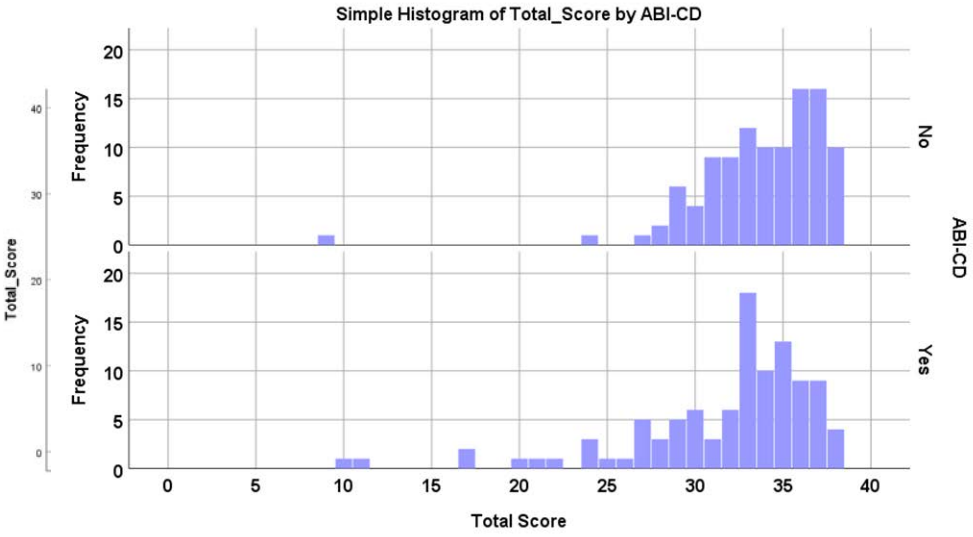
There are two things to consider in looking at data patterns from two groups being compared. The first thing is if the data from groups has the same center point (known as the average). The second thing we want to know is if the distribution of measured values has the same shape. If the center points and shapes are the same, we conclude there are no important differences between compared groups. Groups in the following figures are defined as cases with a self-reported history of ABI (ABI-CD Yes) and cases not reporting ABI
(AB-CD No).

Histograms show the relationship between outcome scores and the number of cases having specific scores. Another way to visualize the distribution of the measured outcome pattern is a boxplot shown below:

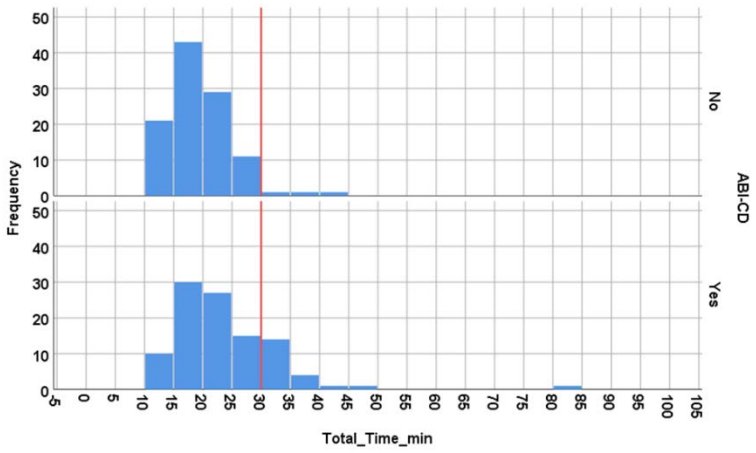


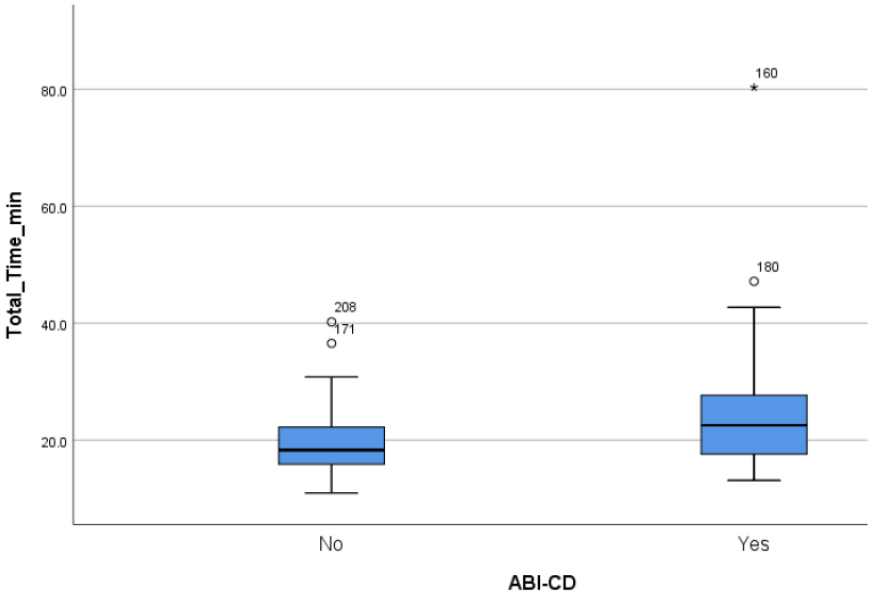
If data distributions from to groups is the same, all the parts of the boxplot should line up, with the values on the left side of the diagram looking the same. Results comparing various outcomes for ABI Yes and No are shown below.

A. Total accuracy scores

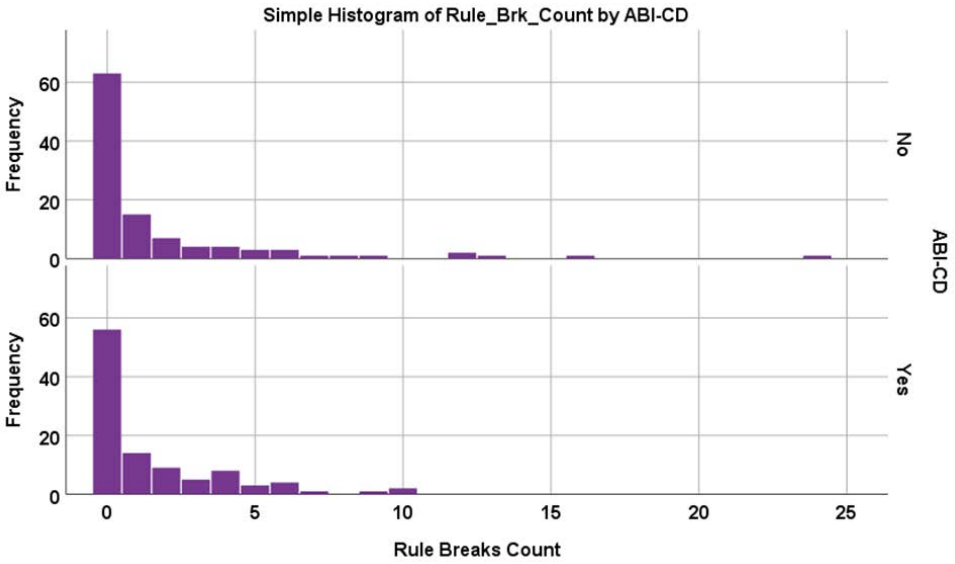


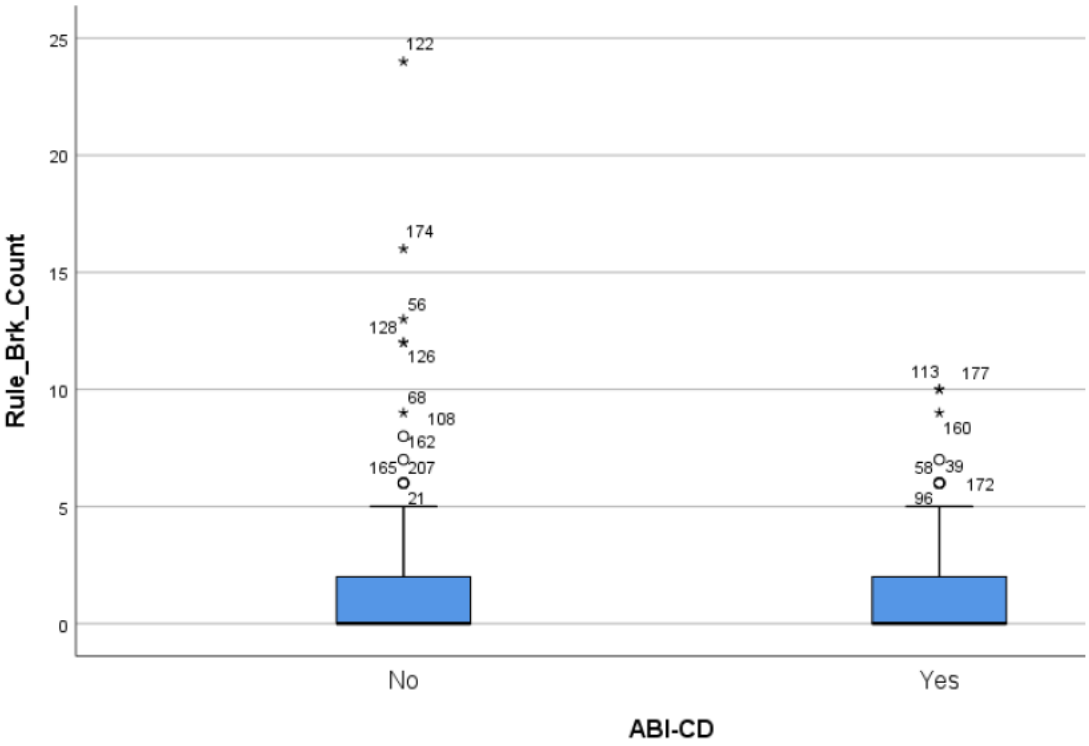
B. Total time



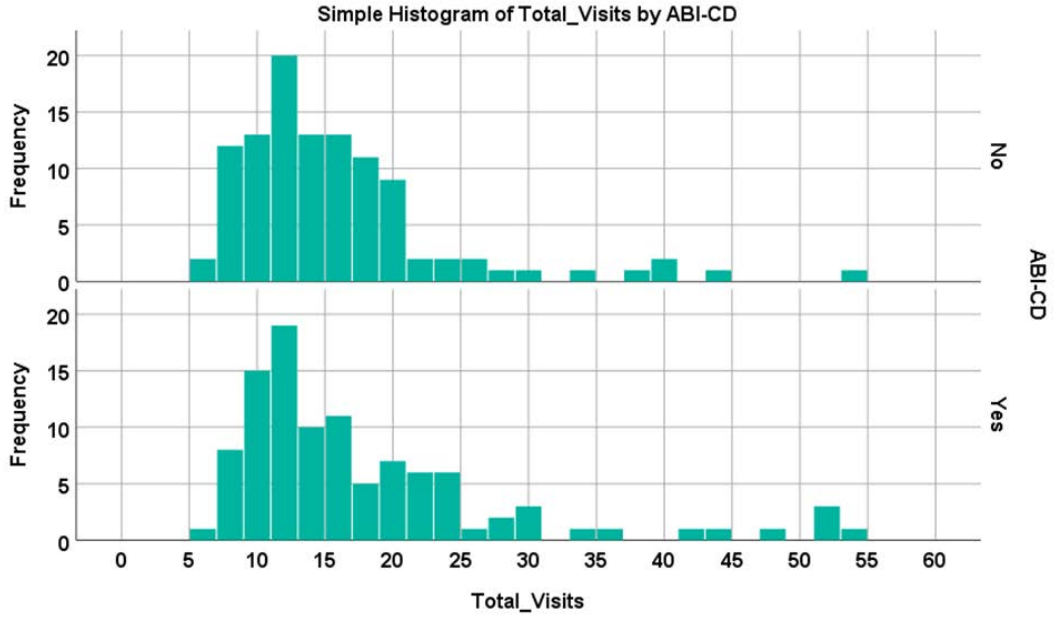


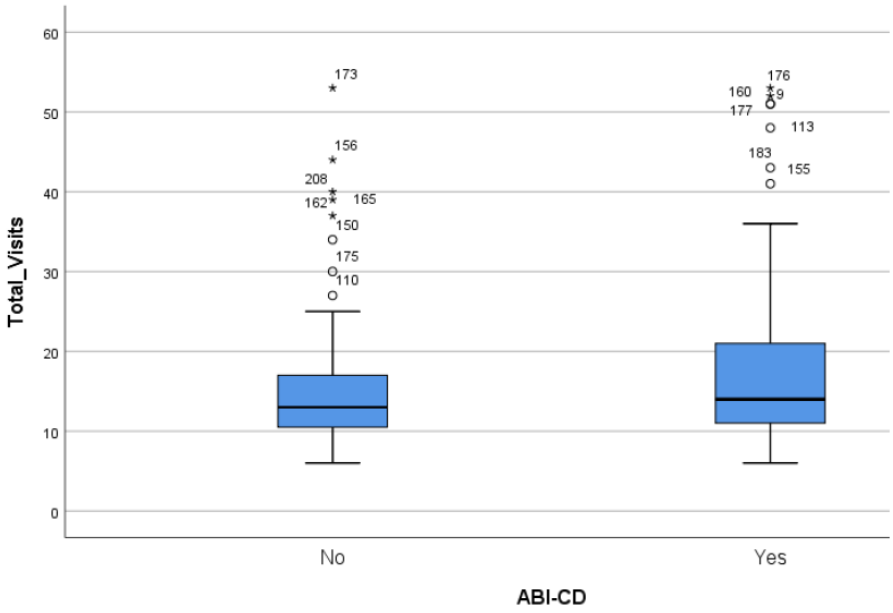
C. Total Rule Breaks



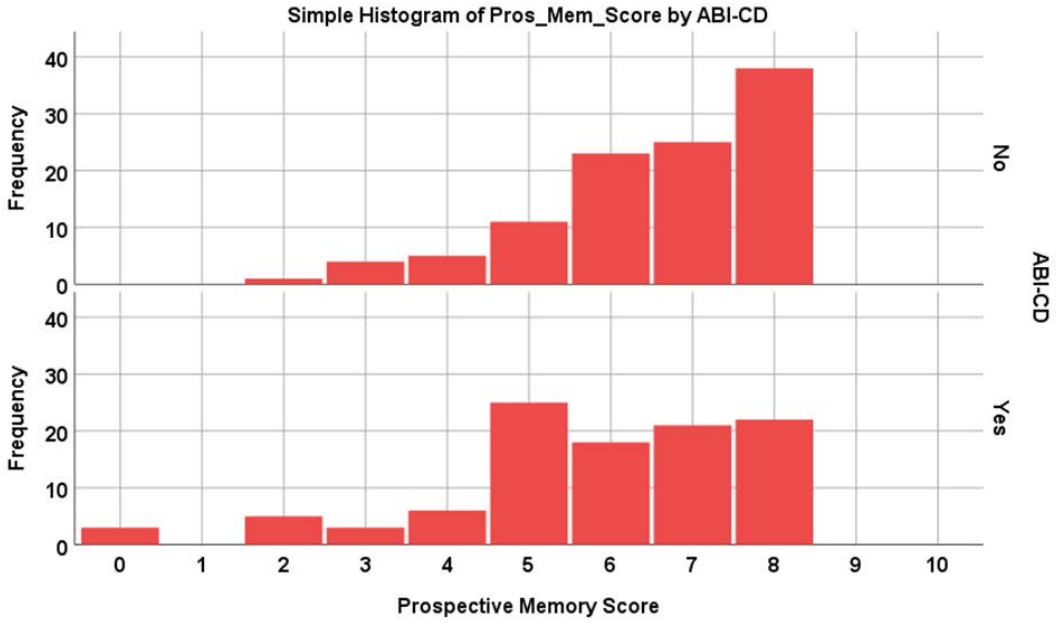


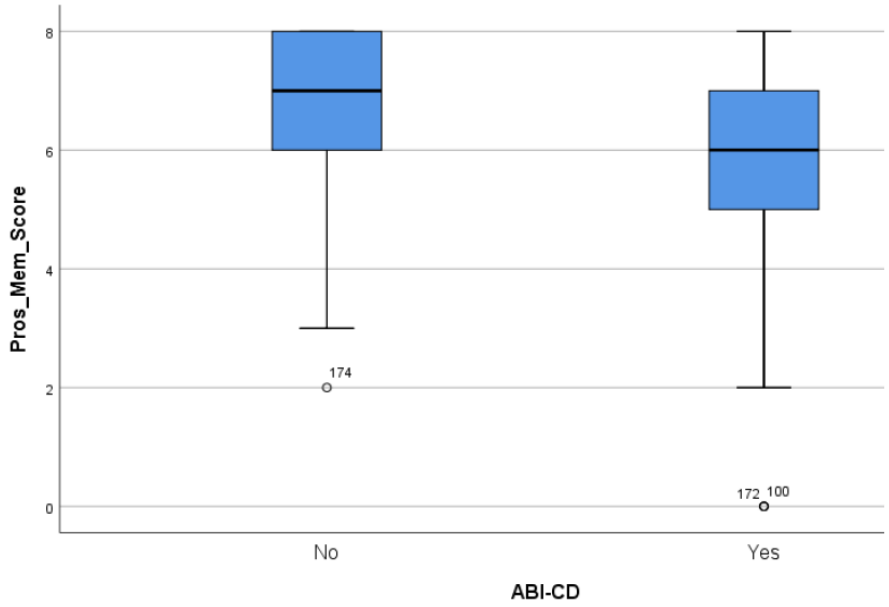
D. Total number of visits





E. Prospective memory items





2.2.2.3 Interpretation

There are no radical differences in data patterns for ABI Yes and No groups. Pattern concordance is not absolute, but the group differences are minor. A useful rule of thumb is if the distribution centers vary by a factor close to 2, or a lot of outliers are seen in one group compared to another, and their range is different. A one-point difference is median can be seen in some cases, but that should be weighed against the clinical significance of this small difference. When the box plots no longer overlap, then it is time to get excited.

The one outcome of marginal interest was for B. Total Time. Note that there is an apparent cutoff in the histogram at 30 minutes, where cases not having ABI had almost no cases above that value, where the ABI group had quite a few. And the tail for ABI is longer, with observations going much further out than non-ABI cases. This is not a large effect, but does indicate some difference in kind between the ABI Yes and ABI No groups.

2.2.3 Might multitasking test score values at the 90th percentile (i.e., the lowest decile of subscores) serve as a useful metric for identifying problematic multitasking test performance?

2.2.3.1 Approach

In order to inform team discussions aimed at identifying sub-scores that were indicative of possible impaired performance, we explored the utility of using sub-scores at the 90th percentile as indicators of problematic performance. The team recommended exploration of scores at the 90th percentile because examination of score distributions suggested that many participants had scores at the 75th percentile, suggesting that those scores were not sufficiently unusual or indicative of possible impaired performance. We had sufficient cases to set the threshold higher, and this also improved the effect signal to noise ratio, making any potential data patterns easier to see.

2.2.3.2 Observations

We conducted analyses using the Total Visits score to evaluate the utility of using scores at the 90th percentile to identify problematic multitasking performance. We selected this particular subscore as the "master predictor" because it represents the primary hypothesized challenge of the test, that being to figure out how to complete the test assignments in as few visits as possible. Based on scores from 210 participants: mean number of visits = 16.40; median number of visits = 14.00; 90th percentile value = 27.

We further evaluated mean subtest scores based on whether or not participants' # of visit fell at or below the 90th percentile (1 group in Table 17). Mean subtest scores were compared based on 2 groups (0=# of visits > 90th percentile; 1 = # of visits ≤ 90th percentile) using the Independent Samples Test; there were statistically significant group differences (p=.000 for each - total visits, rule break count, total time, prospective memory score).

These findings suggested that when a person performs the multitasking test and has a Total Visits score in the 90th percentile, the other sub-scores also likely shift in a negative direction as well. Further, the other subtest scores are significantly different when comparing scores of people whose Total Visits fell in the 90th percentile versus those whose Total Visits fell in the more normal range.

Table 17. Mean multitasking test scores based on whether or not participants' "visits" score fell at or below the 90th percentile

	Visit category	N	Mean	SD	p value for group differences based on visit category
Total performance accuracy score	0	188	33.35	4.284	p=0.000
	1	22	28.55	6.360	
Rule break count	0	188	1.38	2.997	p=0.000
	1	22	3.91	3.146	
Total time in minutes	0	188	20.827	6.0728	p=0.000
	1	22	28.252	14.1147	
Prospective memory items total	0	188	6.36	1.598	p=0.000
	1	22	4.95	2.257	

Table 18. Multitasking subtest cut-off scores at the 90th percentile for "worst" performance (n=210)

Multitasking Subscore	Cut-off value
Performance accuracy (total score)	28 or less points
Total rule break counts	5 or more rule breaks
Total performance time	31 minutes or more to complete
Total number of visits	27 or more visits to complete
Prospective memory subscore	4 or less points

2.2.3.3 Impressions and recommendations for multitasking test score interpretation

These analyses affirm the potential utility of using test score values in the 90th percentile to identify problematic test performance. The research team concluded that if a participant has 2 subscores at the 90th percentile, clinicians should consider further treatment or evaluation for executive dysfunction.

However, performance of complex multitasking tests such as the CQDT, MODT, and FDDT should never be evaluated solely by subtest scores. Clinicians are advised to consider behavioral observations during test performance such as the patient's non-verbal responses starting at the time of presenting the instructions. For example, the patient may take many notes as a memory compensation strategy, they may look very confused or nonchalant and seemingly not paying attention. The clinician can respond to questions from the patient at this time and is encouraged to clarify instructions prior to beginning the test. Other behaviors that a clinician may observe during the test include indicators of the patient's executive functioning skills. For example, planning or lack of planning demonstrated by sitting at the desk for a prolonged time as they create a strategy for completing the tasks versus moving quickly from work area to work area, which may indicate a more impulsive, less planned approach. The clinician is required to fully attend to the non-verbal communication of the patient throughout the test. For example, at some point a patient may reach their threshold for tolerating the demands of the tasks and lose concentration, which can affect their overall performance. If the clinician notices and records an occurrence such as this, they can process the situation with the patient at the conclusion of administering the test. The clinician should pair these behavioral indicators and actions with how they relate to the patient's day-to-day life in order to provide insight into the potential consequences of their demonstrated challenges with multitasking. These clinical observations will contribute to the future treatment provided to the patient by the occupational therapist.

REFERENCES

Carroll, G. L., Radomski, M., Halsten, J., Meyers, A., Rabusch, S., Takimoto, R., & Kreiger, R. (2020). Front Desk Duty Multitasking Test after Mild Stroke: Preliminary Reliability and Validity. *Canadian Journal of Occupational Therapy, 87* (5), 372-381. https://doi.org/10.1177/0008417420953227

Radomski MV, Davidson LF, Smith L, Finkelstein M, Cecchini A, Heaton KJ, McCulloch K, Scherer M, Weightman M M (2018). Towards return to duty decision-making after military mild traumatic brain injury: Preliminary validation of the Charge of Quarters Duty Test. *Military Medicine*, 183(7), e214-e222 DOI: 10.1093/milmed/usx045.

Smith, LB, Radomski, MV, Davidson LF, Finkelstein, M, Weightman, MM, Scherer, MR, McCulloch, K (2014). Development and preliminary reliability of a multitasking assessment following concussion. *American Journal of Occupational Therapy, 68,* 439-443.

Appendix B

Maintenance Office Duty Test (CQDT-Alternate Form)

Administrative Manual

TEST SET UP
Maintenance Office Duty

I. Description: The SM is challenged to develop and execute a work plan for completing an array of interleaving assignments associated with his/her hypothetical assignment to Maintenance Office duty.

II. Purpose: This multi-task provides an opportunity to observe and quantify planning and implementing a plan; specifically, how a SM with mTBI approaches an unstructured complex task when only task parameters and outcomes are specified.

III. mTBI-related task challenges: Primary ● Secondary o

Cognitive				Sensory			Physical			
Executive function	Memory	Attention	Reaction time	Eye gaze tracking	Scanning	Vestibular	Balance	Exertion	Bend-lift	Manual UE Speed
●	o				o				o	o

IV. Source: Adapted from *Multiple Errands Test-Simplified Version* (Alderman et al., 2003)

Alderman, N., Burgess, P.W., Knight, C., & Henman, C. (2003). Ecological validity of a simplified version of the multiple errands test. Journal of the International Neuropsychological Society, 9, 31-44.

V. Materials and Supplies Set up and administration items:

- Blue painters tape
- Tape-measure
- Clipboard
- Administration manual and scoresheet
- Stopwatch
- Pencils

Test items:

Laminated signs to be posted in each work area (Work Station, Stock Room, Break Room, and Maintenance Office)
Wall clock

Work Station –

- Table & Chair

Examiner sits or stands in a location to fully observe the table in order to observe rule adherence regarding number of PVC parts in the Work Station at any given point in time.

Stock Room –

- Bag to carry items
- Duffel bag to hold the following items – o 1" diameter PVC pipe* (4, 12.5" in length; 4, 20.5" in length; 4, 8" in length, 4, 4.5" in length) o T-Connectors (4, 1" diameter T- PVC connectors)
    - o 3-way Connectors (4, 1" diameter 3-way elbow PVC connectors) o 4-way Cross Connector (1, 1" diameter 4-way cross fitting connector)
    - o Measuring Tape
- Fully Stocked Maintenance Department Supplies List sign posted on wall • Table on which to place bag is optional

- Electrical Panel affixed to cardboard with cover so panel is not visible while subject is standing outside Stock Room work area

*PVC pipe and connectors may be purchased at local home improvement stores or ordered directly from Formufit, Inc. http://www.formufit.com/

Break Room –

Signs affixed to cardboard display board: staff duty schedule, parking lot map, diagram for trash can frame assembly, 2 foils (information not relevant to the task) (see Figure for layout). Blue tape must extend for the length of the posted materials and be ~ 4 feet from the wall (to prevent subjects from reading the contents from outside the designated area). **Select a space in which information posted in the Break Room area cannot be read while the subject is standing in any of the other work zones – using an easel for positioning works well.**

Maintenance Office desk – Table
or desk
Chair
2 walkie talkies**
Basket for walkie talkies
Binder with 10 tabbed pocket folders:
Maintenance Duty Reports
Completed Maintenance Office Duty Reports
Inventory Forms
Completed Inventory Forms
Maintenance Tracking
Parking Lot Repair Logs
Staff Duty Schedules
Emergency contacts
2 blank folders
Laminated Contact List (placed in the front pocket of the binder)
Laminated Communications Roster (positioned on desktop)
Laminated Walkie Talkie instructions (positioned on desktop)

**Walkie talkies may be purchased at local home improvement stores. Laminated Walkie Talkie instructions include in the Manual pertain to the Cobra MicroTalk made by Cobra Electronics. Laminated instructions may need to be modified for other models. https://www.cobra.com/products

V. Test Set Up:

Space estimate: Approximately 6 feet by 11 feet area

The Maintenance Office Duty test can be set up in whatever configuration aligns with the available clinical or testing space (Figure 1). Work area boundaries are established by the placement of blue painter's tape. The measurements in Figures 2 – 5 are estimates. Testers should use the checklist below to make sure that tape placement assures that SM must step into the work area in order to complete relevant test tasks.

Place an X outside of the work areas for the starting place.

Figure 1. Example of how the Maintenance Office

Figure 2. Maintenance Office Desk

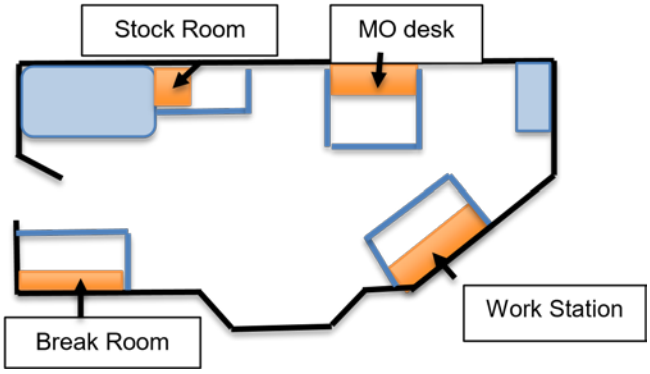


Figure 3. Work Station

Duty test might be set up in a clinical space.

Figure 4. Stock Room

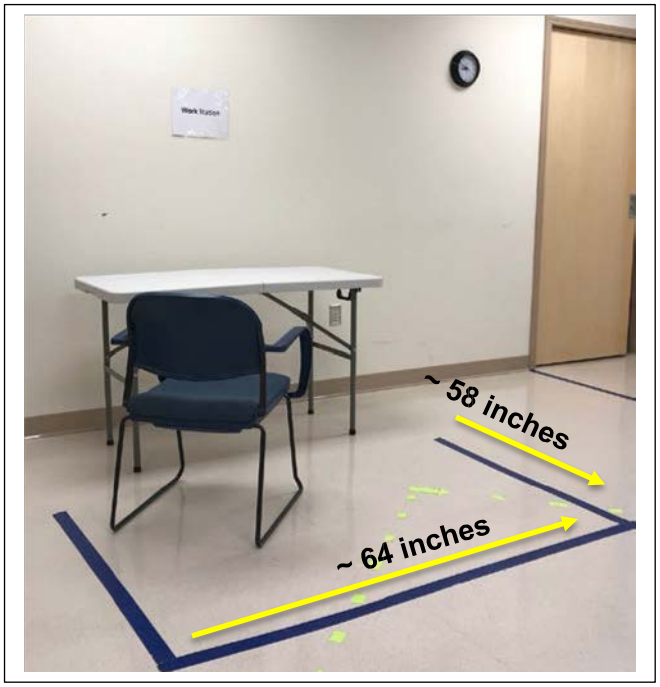
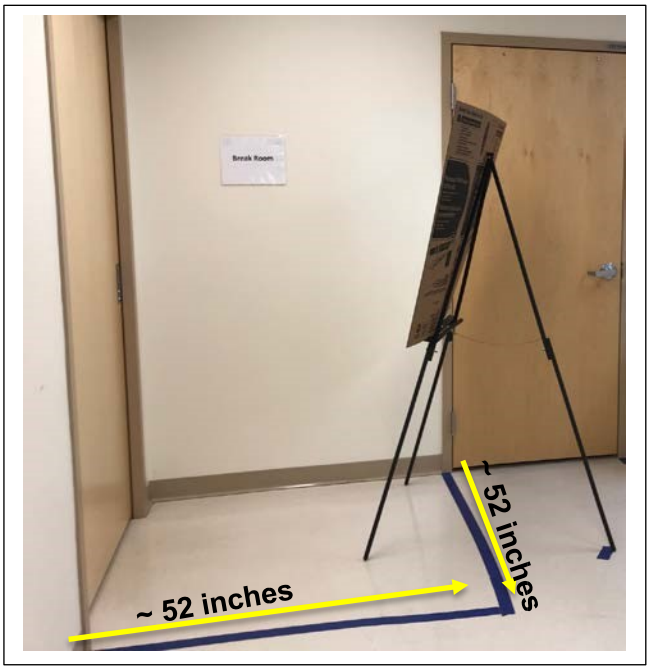


Figure 5. Break Room



Test Set Up Checklist:

____ Is the wall clock visible from all work areas?

____ Are the Staff Duty Schedule, Parking Lot Map and Trash Can Frame diagram visible ONLY if standing inside the blue tape in the Break Room work area? [If one or more can be read from outside the blue tape, move the tape further out or use an easel for positioning.]

____ Is it impossible to view the trash can frame diagram in the Break Room from the Work Station? [Select a location for the Break Room Board in which SM cannot see the diagram from the Work Station. Consider using an easel and affixing signs to cardboard.]

____ Is the Contact List and Communications Roster readable ONLY if standing inside the blue tape of the Maintenance Office Duty Desk work area? [If one or more can be read from outside the blue tape, move the tape further out.]

Maintenance Office Duty Test (MODT)
EXAMINER INSTRUCTIONS AND SCRIPT

Before instructing the participant, turn on participant's walkie talkie and set to radio frequency 6 and place on the desk in the Maintenance Office; turn on the examiner's walkie talkie and set to radio frequency 2. Place the Communications Roster on the desktop. Make sure a clock can be seen from all areas. During the test, sit or stand in a location that allows you to fully and easily observe the table in the Work Station (WS) in order to observe rule adherence regarding # of parts in the WS at any given point in time.

INTRODUCTION
Provide participant with Maintenance Office Duty Report affixed to clipboard and state the following:

During this test, you must complete some multitasking assignments while on duty in the Maintenance Office. The test is designed to be challenging so that no one can do it perfectly. (Give clipboard and Maintenance Office Duty Report to the participant.) **This is your copy of the Maintenance Office Duty Report. It includes instructions, your assignments, and rules. You can use it throughout the test along with whatever devices and techniques that you think will help you perform at your best. I will orient you to your assignments, the test spaces and the rules. You may ask me any questions before the test begins.**

INSTRUCTIONS
A. Description of assignments

Follow along on your report as I describe your assignments. During the test you should do the following:

- **Radio the Staff Duty Officer at headquarters and state you are reporting for duty as assigned.**
- **Get the current maintenance tracking sheet from the Maintenance Office • Assemble a trash can frame for the Maintenance Office • Before you complete the test:**
  - **Return all supplies and materials to their original locations** o **Place the trash can frame in the Maintenance Office**
- **File your completed Report in the Completed Maintenance Office Duty Reports binder pocket in the Maintenance Office.**
- **Radio the Staff Duty Officer when you have completed this test • Exactly 5 minutes after you start the test:**
  - **Radio First Staff Sargent (1SG) and report the number of vehicles that need refueling.**

- **Also, ask about any additional assignments that you are to complete while on duty.**

- **On your Maintenance Office Duty Report, write down the following 4 items:**
  a) **What is the number of the electrical circuit that needs to be shut off for parking lot repairs (SR)?**
  b) **Record the name and phone number of the base Electrical Company (SR).**
  c) **Record the radio frequency for the guard shack at the entrance of the parking lot.**
  d) **Record the telephone number for the service member on duty next Monday.**

• Write legibly. You will not receive full points if I cannot read your handwriting.

Next, I am going to point out your Work Areas for this test. But first, do you have any general questions about your assignment so far? If yes, clarify. If no, begin orientation to work areas.

B. Orient participant to the Work Areas.

I will now orient you to the Work Areas. Everything you need to complete this test is located in the 4 work areas marked with blue tape.

B1. Walk to each of the areas with the participant to point out location of materials and area perimeter. Start by walking to the Maintenance Office. (MO)

- **We are currently in the Maintenance Office (MO).**
- **Use this walkie talkie during the test. Hold this button down** (point) **while speaking into the walkie talkie; To change the radio frequency, press this button** (point to mode button) **until the channel number flashes, then press the arrow on the right side up or down to find the desired radio frequency. Use these instructions if you need more help using the walkie talkie** (point).
- **Use this Communication Roster** (point) **to locate the correct radio frequency channels for various personnel that you need to contact during the test.**
- **Obtain and file required forms here** (point to binder). **Also note the pad of paper and pen/pencil and a Contact List** (point to list in the inside cover of the binder).

B2. Walk over to Break Room. (BR)

- **Here is the Break Room. Here you will find information that you need to report to command, including a staff duty schedule, a map of the base parking lots areas, and a diagram regarding how to assemble the trash can frame. Note that this diagram is the only information you have about how to assemble the trash can frame.**

B3. Walk to the Stock Room. (SR)

- **This area is the Stock Room. All the materials you will need for assembling the trash can frame are located in this duffle bag. Here are the inventory cards** (point) **that specify how many items we currently have on hand. The electrical panel is also located in this area** (point)**. Additionally, here is a crate that you may use to carry items from one area to another.**

B4. Walk to the Work Station. (WS)

- **We are now in the Work Station. Assemble the trash can frame only in this area.**

Do you have any questions about the work spaces? (If yes, clarify. If no, continue.)

C. Review the test rules.

I will now explain the test rules. During this test, you must follow these rules:
(point out rules on the bottom of the Maintenance Office Duty Report.)

- **You should carry out all of these assignments but may do so in any order.**
- **Assemble the trash can frame only in the Work Station.**
- **Bring only the number of parts needed for the trash can frame from the Stock Room to the Work Station.**
- **Do not move or remove any of the signs or instructions from the work areas during the test.**
- **In order to score the most points, your trips between work areas should be kept to a minimum. You should return to an area only if it is absolutely necessary to complete the assignment correctly. You will get the most points if you can complete the test in 7 trips.**
- **Do these assignments as quickly as possible without rushing.**
- **Do not ask questions for further guidance about this test after it begins.**

Do you understand these rules? If no, clarify. If yes, continue.

EXAMINER GUIDANCE:

If the participant asks whether he/she can use his alarm, phone, watch during the test, state: **"Use any strategy or device that you think will help you do your best."**

If the participant asks any questions about the test and/or procedures BEFORE the test begins, answer the questions directly.

D. After verifying that participant understands the instructions.

Move outside of the designated work areas into a neutral zone. Using the Maintenance Office Duty Report, the participant restates the test instructions and rules.

Now brief me on what you are being asked to do. Fill in any gaps that the participant may not have included.

After doing so and answering all questions about the test, place the participant on the start X outside of all taped areas. State: **Remember, you must complete all of the assignments but may do so in any order.** (When you are ready) **Start.** (begin the timer)

[PARTICIPANT BEGINS THE TEST.]

EXAMINER GUIDANCE: IF THE PARTICIPANT IS USING THE WRONG RF

If the participant tries to perform step 1 (radioing Staff Duty Officer) but is not on the correct frequency, state: **I can't hear you on the radio.**

If the participant still does not figure out to change the rf, state: **you are on the wrong rf.** See side 2 of score sheet for scoring instructions.

EXAMINER GUIDANCE: HOW TO RESPOND TO QUESTIONS ASKED DURING TESTING AND SCORE RULE BREAKS RE: QUESTIONS

- Scoring: Every question asked = 1 rule break.
- If a participant asks a question during the test, state: **"Do what you think is best."**
- If the participant has not figured out a solution to his/her own question **within 1 minute**, point to the relevant information on his or her Maintenance Office Duty Report and provide information to get him or her back on track. [Scoring: The item for which the participant needs this examiner assistance is scored a 0.]

E. Re-set examiner walkie talkie during test.

At the beginning of the test, the examiner's walkie talkie is set at frequency 2, change to frequency 5 for five minute call in.

EXAMINER GUIDANCE: WALKIE TALKIE RESPONSE TO REPORT TO STAFF DUTY OFFICER

Participant: "Staff Duty Officer this is XXX, Over"
Guard Shack: **"Xxx this is SSC Smith, go ahead over"**
Participant: "Beginning Maintenance duty"
Guard Shack: **"Good copy, over." Or "Roger that, Over."** Participant: "Over"

After the subject radios to check in (**assignment 1** on score sheet), the examiner changes the radio to frequency 5 (corresponding to 1SG and **assignment 4** on score sheet). **[5 MINUTES INTO TEST]**

F. Provide additional instructions when participant radios in after 5 minutes and asks about more assignments.

- **Get an inventory form from the Maintenance Office binder**

- **In the Stock Room, inventory the supplies that need to be re-ordered using the Current Inventory Cards and the Fully Stocked Supply sheet found in that area.**
- **When done, place the completed form in the completed inventory forms binder pocket in the Maintenance Office.**

EXAMINER GUIDANCE: WALKIE TALKIE DIALOGUE WITH

Participant: **"First Staff Sergeant this is XXX, Over"** First
Staff Sergeant: This is First Staff Sergeant. Over.

- Participant: Reporting number of vehicles in **section A** for refueling. Any additional assignments over?

First Staff Sergeant:

- **Get an inventory form from the maintenance desk**
- **In the Stock Room, inventory the supplies that need to be re-ordered using the Current Inventory Cards and the Fully Stocked Supply sheet found in that area.**
- **When done, place the completed form in the completed inventory forms folder in the Maintenance Office.**

Participant: Repeats instructions, clarification made if error. If correct, "Good copy, Over"

NOTE: If the subject does not stop and radio in for further instructions by 6 minutes and 30 seconds into the test, state:

Stop and do what I instructed you to do 5 minutes into the test.

After the participant completes **assignment 4,** the examiner resets the walkie talkie to frequency 2 (corresponding to assignment 13 on score sheet).

G. The test is discontinued when:

a) The Participant radios report to the Staff Duty Officer that he/she has finished thetest OR;
b) The Participant states he/she does not want to continue OR;
c) Participant demonstrates behavioral contraindications as specified in IRB application OR;
d) The Participant is still performing the test at 30 minutes and does not appear to bewithin 2 minutes of completion.

Maintenance Office Duty Report

Study ID #: _____

You should do the following:

- Radio the Staff Duty Officer at headquarters and state that you are reporting for duty as assigned.
- Get the current Maintenance Tracking Sheet from the Maintenance Office.
- Assemble a Trash Can Frame for the Maintenance Office.
- Before completing the test o return all supplies and materials to their original locations o place the Trash Can Frame in the Maintenance Office
- File your completed report in the "Completed Maintenance Duty Reports" binder pocket in the Maintenance Office.
- Radio the Staff Duty Officer when you have completed this test.

Exactly 5 minutes after you start the test:

- Radio First Staff Sargent (1SG) and report the number of vehicles that need refueling.
- Ask about any additional assignments that you are to complete while on duty.

You should obtain the following information during the test: *(write legibly)*

1. What is the number of the electrical circuit that needs to be shut off for parking lot repairs?

   _____

2. Record the name and phone number of the base electrical company:

   _____

3. What is the radio frequency for the Guard Shack at the entrance of the parking lot? _____

4. What is the telephone number for the Service Member on duty next Monday? _____

Rules:

- You should carry out all these assignments but may do so in any order.
- Assemble the Trash Can Frame only in the Work Station.
- Bring only the number of parts needed for the Trash Can Frame from the Stock Room to the Work Station.
- Do not move or remove any signs or instructions from the work areas during the test.
- In order to score the most points, your trips between work areas **should be kept to a minimum**. You should return to an area only if it absolutely necessary to complete the assignment correctly. You will get the most points if you can complete the test in 7 trips.
- Do these assignments as quickly as possible without rushing.
- Do not ask questions for further guidance about this test after it begins.

Participant ID:	Rater:	Date:	Session #:
Did SM complete test? _____Yes _____No (Examiner stopped) _____No (SM requested to stop)			

Maintenance Office Duty Score Sheet

Assignments 2 = 100% accurate, no cues required 1 = Completed but only partly correct OR *1 = Required cues (1,3, 4, 6, 13 only*)* 0 = Did not complete or perform	Score 0, 1, 2	Sub-totals	Scratch Pad for visits to WORK AREAs
1- Radio Staff Duty Officer at headquarters and….[rf 2] STATE*: ***I can't hear you on the radio.***			MO, SR, BR, WS Start (outside work areas) … 1.To _____ 2.To _____ 3.To _____ 4.To _____ 5.To_____ 6.To _____ 7.To _____ 8.To_____ 9.To _____ 10.To _____ 11.To _____ 12.To _____ 13.To _____ 14.To_____ 15.To _____ 16.To _____ 17.To _____ 18.To _____ 19.To_____ 20.To_____ 21. To_____ 22. To_____ 23. To_____ 24. To_____ 25. To_____ 26. To_____ 27. To_____ 28. To_____ 29. To_____ 30. To_____ 31. To_____ 32. To_____ 33. To_____ 34. To_____ 35. To_____
2- Report that you are taking over duty as assigned.			
3- Between 4 min 30 sec -5 min 30 sec, participant stops what he/she is doing* ….NOTE: IF THE SM DOES NOT STOP WHAT HE/SHE IS DOING BY 6 MINUTES & 30 SEC, STATE*: ***Stop and do what I instructed you to do 5 minutes into the exercise.***			
4- Radio 1SG* [rf 5] CUE: *I can't hear you on the radio.* NOTE: IF THE SM DOES NOT KNOW WHAT TO DO AT THE 5 MINUTE INTERVAL, STATE*: ***Check your Maintenance Duty report to see what you are supposed to do now.***			
5- Report # of vehicles that need refueling. [3]		/10	
6- Ask if there are any additional assignments to be completed*. NOTE: IF THE SM DOES NOT ASK THE QUESTION, STATE*: ***What else are you supposed to ask me?***			
7- Get an Inventory Form at the Maintenance Desk			
8- Assemble the Trash Can Frame			
9- Put the Trash Can Frame in the Maintenance Office before completing the exercise.		/8	
10- *Complete the Inventory form based on supply status on supply cards in the stock room.*			
11- *Files Form in Completed Inventory binder pocket*			
12- Return all supplies and materials to their original locations before completing the exercise.			
13- Radio Staff Duty Officer…. [rf 2] *STATE*: I can't hear you on the radio.*		/10	
14- Report test completion.			
15- Write down number of the electrical circuit that needs to be shut off for parking lot repairs. [15]			
16- Record the name and telephone # of the base Electrical Company. [Fort Campbell Circuit Breakers 270-798-5425]			
17- Record the rf for the Guard Shack at the entrance of the parking lot. [10]			
18- Record the telephone # of SM on duty next Monday. [212-756-4594]			
19- File completed Maintenance Duty Report in "Completed Maintenance Duty Reports" binder pocket		/10	
Assignment performance total score	**A.**_____/38		
Rules (see rule break definitions on back)	**Rule break (Y/N)**	**# of rule breaks**	
Do not ask questions for further guidance about assignments once the test starts.			
Assemble the Trash Can Frame only in the Work Station.			
Bring only the number of PVC parts needed for the Trash Can Frame to the Work Station [25 parts].			
Do not move or remove any signs or instructions from the work areas.			
Everything else must be scored in real time EXCEPT 10,11,15-19	**B.**	**C.**	**D.** Total # of visits to complete exercise: _____
Performance time: **E.** _____min _____sec	**Visit** = any body part crosses into taped work area.		

Rules	Rule break examples
Do not ask questions for further guidance about assignments once the test starts.	1 rule break for **every question** asked. Making statements aloud ≠ rule break. ["Can you help me get this walkie talkie to work? What is the correct RF?" = 2 rule breaks]
Assemble the Trash Can Frame only in WS.	Each time SM puts 2 parts together outside of WS = 1 rule break SM connects PVC T connector to 6" piece in SR = 1 rule break
Bring only the # of parts that are needed for the Trash Can Frame to the WS [25]	If SM has 23 parts at WS = 2 rule breaks SM brings tape measure to WS = 0 rule break (not a needed part)

Tasks	Scoring examples (1, 0)
1. Radio Staff Duty Officer and.. [rf 2]	1 = radios Staff Duty Officer after cue re rf 0 = does not do this assignment at all **OR** requires further cueing re rf
2. Reports taking over duty as assigned	1 = reports something other than that he/she is taking over duty 0 = does not do this assignment at all
3. Between 4 min 30 sec -5 min 30 sec, participant stops what he/she is doing ….	1 = stops what he/she is doing between 5 min 30 seconds and 6 min 30 seconds 0 = called before 4 min 30 sec **OR** didn't radio in by 6 min 30 sec **OR** examiner instructs to stop what he/she is doing now
4. Radios 1SG [rf 5]	1 = radios 1 SG after cue re rf 1 = radios after being cued by examiner to check Maintenance Duty Report 0 = even with cue, SM doesn't know to radio 1SG and is instructed to do so
5. Report # of vehicles that need refueling. [3]	1 = reports incorrect # of vehicles 0 = doesn't report this information at all
6. Ask if there are any additional assignments to be completed while on duty*.	1 = asks if there are additional assignments to be completed after being cued 0 = even with cue, SM does not ask the question and the examiner simply provides the 3 additional assignments (7,10, 11)
7. Obtain an inventory form from the files at maintenance desk	1 = obtains the wrong form from maintenance desk 0 = doesn't get any form from maintenance desk
8 . Assemble a Trash Can Frame	1 = constructs Trash Can Frame, but made errors related to 1 – 2 parts 0 = does not do this assignment at all **OR** attempted with errors on 3 or more parts
9. Put the Trash Can Frame in the MO before completing the exercise.	1 = puts the Trash Can Frame on the desk 0 = puts Trash Can Frame in another work area
10. *Complete the inventory form*	1= takes inventory but some of the values are incorrect or missing or not legible 0 = does not take inventory
11. *Files inventory form in "Completed Inventory Form" folder*	1 = files the form but in the wrong folder **OR** form in binder but not in folder 0 = does not file the form
12. Returns all supplies and materials to their original locations before completing the exercise.	1 = returns some but not all supplies and materials to original locations 0 = does not return any of the supplies and materials to their original locations before radioing that he/she is done with maintenance duty
13. Radio Staff Duty Officer…. [rf 2]	1 = radios Staff Duty Officer after cue re rf 0 = does not do this assignment at all **OR** requires further cueing re rf
14. …to report exercise completion.	1 = reports something other than test completion **OR** calls in <u>before</u> completing test (i.e., filing maintenance duty report as instructed) 0 = does not do this assignment at all

15. Write down number of the circuit that needs to be shut off for parking lot repairs. [15]	1 = fills in this part of maintenance duty report but with wrong circuit #
	1 = identifies circuit # but not clearly legible
	0 = does not fill in this part of maintenance duty report

16. Record the unit name and phone number for the base's Electrical Unit. [Fort Campbell Circuit Breakers 270-798-5425]	1 = fills in only a portion of the information requested or makes errors in the information that was written (includes misspellings, transposes digits, etc.) 1 = completes information but some or all is not legible 0 = does not fill in this part of the maintenance duty report
17. Radio Frequency for the guard shack at the entrance to the parking lot [10]	1 = fills in but makes errors in the rf that was written (transposes digits, etc.) 1 = completes this part, but some or all is not legible 0 = does not fill in this part of the maintenance duty report
18. Record the telephone # for the service member on duty next Monday. [212-756-4594]	1 = fills in name of Michaels but not phone # 1 = fills in wrong phone number 1 = fills out form but answers are not legible 0 = does not fill in this part of maintenance duty report
19. File completed Maintenance Duty Report in the "Completed Maintenance Duty Reports" folder	1 = files the form, but in the wrong folder **OR** form in binder, but not in folder 0 = does not file the form

Test performance is over immediately after the SM radios Staff Duty Officer & reports test completion. Stop timer and record the performance time. **SM scores a "0" for any assignment completed <u>after</u> radioing Staff Duty Officer to report test completion.**