A Functional Assessment of Executive Functioning: The Hamburger Turning Task

by

Sara L. Shugars

Psychology, BS, Michigan State University

Submitted to the Graduate Faculty of

The School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Master in Science

University of Pittsburgh

2007

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTH AND REHABILITATION SCIENCES

This thesis was presented

by

Sara L. Shugars

It was defended on

April 5, 2007

and approved by

Michael Pramuka, Ph.D., Assistant Professor

Shirley Fitzgerald, Ph.D., Assistant Professor

Thesis Advisor: Michael McCue, Ph.D., Associate Professor

A FUNCTIONAL ASSESSMENT OF EXECUTIVE FUNCTIONING: THE HAMBURGER TURNING TASK

Sara L. Shugars, M.S.

University of Pittsburgh, 2007

The number of neuropsychological tests and functional assessments that claim to have a relationship between the patient's testing performance and behavior in real-world settings is limited. Additionally, there is a growing concern among testing professionals that most, if not all, psychological tests and standardized assessments introduce environments and stimuli that people never encounter and tasks that often do not emulate life situations or vocational requirements. In order to address the current issues surrounding the ecological validity of psychometric tests, this pilot study introduced a hands-on assessment using a simulated real-world vocational task. Twenty-three subjects between the ages of 18-26, with varying cognitive disabilities, completed a vocational simulation task, the Hamburger Turning Task (HTT), and the results were compared to their scores on a battery of commonly used neuropsychological tests (Wisconsin Card Sorting Test, Finger Tapping test, Controlled Oral Word Association Test, Trailmaking Test, Stroop Color-Word test) that purport to measure aspects of executive functioning analogous to those measured by the HTT. A Pearson-product correlation was run to compare the relationship between the scores from the HTT and the psychometric tests, as well as the relationship between both psychometric tests and HTT scores and daily behavioral observations of executive functioning related performance over a 7 week period. The results of the study found a

significant correlation between the HTT and behavioral data, leading us to believe that the HTT can be used to evaluate real-world aspects of executive functioning. It was also found that there was a high level of interrater reliability on the scoring of the HTT, allowing future researchers to use this as a standardized tool.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

Acknowledgements- I would like to thank each of the members on my committee – Dr. Michael McCue, Dr. Shirley Fitzgerald and Dr. Michael Pramuka- for their time, guidance and insight. In addition, I would like to acknowledge Shannon Juengst, Jon Fienburgh and Sherrol Shugars for all their editing assistance, much needed feedback and around the clock answers to my countless questions.

1.0 BACKGROUND

Originally, neuropsychological tests (NP) were developed to detect cognitive deficits originating from brain injury and to gauge levels of impairment. NP tests are used to evaluate a variety of cognitive abilities, such as: concentration, memory, reading comprehension, judgment, visual perception, movement, coordination, interpretation and processing information. The level of ecological validity, or real-world relevance, of NP testing has been frequently been called into question. The majority of these tests were not designed to predict how individuals are likely to function in real-life settings, live independently and/or remain employed (Sbordone, 2001). The number of tests that measure executive functioning (EF) that are also ecologically valid are difficult to find. This is cause for concern due to the impact that EF has on an individual's ability to function independently on a job and at school. EF encompasses a range of cognitive attributes, including planning, checking, monitoring, testing, evaluating and revising. The purpose of this study is to determine how well a hands-on functional assessment correlates to real-world functioning and to create standardized scoring for this measure, allowing it to be used as a functional assessment with real-world relevance in future research of EF.

This paper presents a discussion of the assessment of EF, a discussion of the ecological validity of NP tests to measure real-world abilities and a review of literature relevant to the measurement of ecological validity of EF measurement.

1.1 EXECUTIVE FUNCTIONING

Executive functions were the focus of this pilot study, which are a highly debated cluster of brain operations primarily associated with the pre-frontal cortex. The executive functions have an impact on cognitive and emotional functioning, more specifically cognitive inhibition, initiation, self regulation and motor output, all of which greatly impact rehabilitation plans and vocational goals. Executive functioning (EF) was initially researched through the observation of individuals with frontal lobe damage (Luria, 1966). Subsequent research and EF assessments were established using populations with similar brain injuries, causing these tests to lack generalizability to other cognitive disabilities. "It seems that standard neuropsychological tests of executive functioning are deemed to be such because of their sensitivity to frontal lobe damage, rather than because they have been operationalized to assess the theoretical concept of EF and the group of cognitive processes it entails" (Bryan & Luszcz, 2000, p.41). The complex nature of executive functions has resulted in differences in its definition, as well as debate on which assessments are appropriate measures for these aspects of brain functioning.

Additionally, researchers have found it challenging to establish how different executive tests assess the different facets of EF, and to differentiate between EF tasks and non-EF tasks (Rabbit, 1997). According to Wong (1985), the specific executive functions include planning, checking, monitoring, testing, evaluating and revising. Thus, they require the ability to mobilize, allocate, coordinate and problem solve. For the purpose of this thesis, "The executive functions can be conceptualized as having four components: (1) volition; (2) planning; (3) purposive; and (4) effective performance. All are necessary for appropriate, socially responsible, and effectively self serving adult conduct" (Lezak, 2004, p. 611). Despite the discrepancies of opinion between

psychologist's definitions and assessments of executive function, the importance of the role it plays in daily functioning and complexity are readily agreed upon by clinicians.

From a clinical perspective, there is a shared understanding that executive functioning plays a substantial part in an individual's ability to operate independently at home, at school or in vocational settings (Cicerone & Giacino, 1992). Considering the significance of the purpose that executive functioning (EF) plays, it is important to be cognizant of the specific ways that dysfunction in this area can impact behavior and of the types of disabilities/diagnoses in which this dysfunction is typically observed. Problems in executive functioning are associated with numerous disabilities, including Attention Deficit Disorder, brain injury, Obsessive Compulsive Disorder, Autism spectrum disorder, Depression, and learning disabilities. Research in EF has primarily concentrated on traumatic brain injury and Attention Deficit Disorder. Although these populations are easily recruited for research, the narrow scope of disabilities studied has left a significant gap in information about other diagnoses.

Researchers have found that it is unusual to encounter an individual with impairment in just one aspect of EF, but rather deficits in EF typically involve a distinctive set of activity related disorders (Lezak et al, 2004). Although NP testing is able to pick up on many cognitive dysfunctions quite often subtle impairments may not appear in the scores (Lezak, 1982). There are a number of standard tests frequently used in clinical settings which are designed to measure different aspects of EF, including the Stroop, Tower of London (Shallice, 1982), Wisconsin Card Sorting Test (WCST; Heaton, 1981), Trail Making Test (Parts A and B) (Reitan, 1958), and the Rey Complex Figure (Lezak, 1983). Lezak (2004) addresses the issues surrounding EF and the typical NP testing constraints in the statement: "A major obstacle to examining executive functioning is the paradoxical need to structure a situation in which the patients can show

whether and how well they can make structure for themselves". Although these tests are considered to be reliable measures and are widely used, there is still a debate on whether or not these clinically administered tests are able to detect the real-life functioning level of the subject.

1.2 ECOLOGICAL VALIDITY

The relationship between the scores or level of performance on neuropsychological tests and an individual's present or future functioning in real-world settings is often referred to as ecological validity. According to Franzen and Wilhelm (1996), two approaches to addressing the problem of ecological validity of assessments and tests have been identified: verisimilitude and veridicality. Verisimilitude includes the inspection of an instrument on the theoretical relationship between the cognitive demands of the test and the situational behavior that the instrument was designed to measure; also how well the instrument matches the demands that are required in real-life. Veridicality is another approach to ecological validity that refers to the extent to which existing tests are empirically related to measures of real-world functioning (Frazen & Wilhelm, 1996). According to Sbordone (1996, p. 23), the level of an instrument's ecological validity was originally based on two main outcomes: ability to return to work and competency in activities of daily living.

The primary goal of testing/assessment is for the clinician to determine how an injury or impairment has affected the individual's ability to live independently, return to work or school, and manage their lives. Most NP tests are not designed to make such determinations, and the ability of the test to do so has been brought into question (Sbordone, 1996). Consequently, there

is a growing concern amongst testing professionals that most, if not all, psychological tests and standardized assessments introduce environments and stimuli that people never encounter and tasks that often do not emulate life situations or vocational requirements.

Ecological validity applies to all neuropsychological tests. For this study, we have focused on the ecological validity of tests and assessments of executive functioning. In the past decade, there has been an increase in the number of tests that have attempted to increase ecological validity in the measurement of executive functioning. "The study of executive functions and their effects can be best understood by a complex evaluation of processes that incorporate objective quantitative as well as objective qualitative methods of observation" (Cripe, 1996, p. 194). There is a trend in neuropsychology toward the development of tests which introduce a variety of new concepts to testing, as well as a concentration on decreasing the tests' structure and increasing their vocational relevance. Lezak (1982) introduced the Tinkertoy Test, which was one of the first tests that attempted to reduce the amount of structure. This test allows the subject to exhibit his/her ability to initiate, plan and structure an activity. Another test created to be more ecologically sensitive, in a less clinical approach, is The Behavioral Assessment for Vocational Skills (BAVS): Wheelbarrow Test (Butler & Anderson et al., 1989), which provides a job simulation task. This environment offers very little structure and presents an opportunity for distractibility, problem solving, impaired judgment and emotional control to be observed. The Multiple Errands Test (MET; Shallice & Burgess, 1991) allows the clinician to see how an individual reacts to new environments, adjusts to real-life settings (shopping mall), and is able to spontaneously initiate strategies. Simulation approaches of testing, like these, present an opportunity for the individuals to organize tasks conceptually in order to continue

with problem solving; these are essential components of EF needed to achieve success in vocational and educational goals.

These tests that claim to have increased ecological validity in testing executive functioning are a direct result of the need for better diagnostic and functional measures, but even these assessments are still limited. New NP tests attempt to address and incorporate real life demands and real life outcomes and are validated by using observations, but they typically measure only a small number of specific executive functions using a limited range of activities. According to Sbordone (1996), the testing/assessment environment is not the real world and is frequently structured in a way that falsely optimizes performance. Often, an impaired individual will function at a level that is much higher or significantly lower than the levels suggested by psychological testing results. The clinical testing environment is ideally designed to be free from distractions and highly structured, and the examiner instructs the subject what to do or not do; feedback is clear and the time demands are stated. "The heavy use of quantitative procedures (test scores) with the minimum use of qualitative procedures (direct observations of the processes) greatly limits the observation of [executive functions]" (Cripe, 1996, p. 194). To get a comprehensive picture of how an individual truly functions in the world, it is imperative that there be both standardized clinical testing as well as functional or situational assessments to provide an ecologically valid facet to the process.

1.3 EXECUTIVE FUNCTIONING, ECOLOGICAL VALIDITY AND HTT

The hamburger turning task (HTT) was developed from training materials for an actual vocational task at a popular fast food establishment. The task requires the individual to move objects (faux hamburgers) in a sequential manner from memory (using spatula and salt shaker), adhere to specific rules, and maintain attention. In the original training task, the participant watched a video simulation of the task, practiced, and then completed the task while being evaluated by the employer. The HTT modifications followed and adapted the instructions that were provided by the staff and training video supplied by the fast food chain. By modifying and using this procedure as a functional assessment tool, we were able to explore how the participant's score on this hands-on task related to his/her performance on structured neuropsychological tests that have established validity, and how this performance related to observation of executive behaviors throughout daily actions in a documented training program.

Shallice and Burgess (1991) noted that in the clinical administration of a test, it is abnormal for a patient to be required to organize or plan behavior over long periods of time, or to set priorities in two or more competing tasks. However, these are the types of executive functions that are a large component of everyday activities (Shallice & Burgess, 1991). A functional assessment is broadly defined as "a measurement of an individual's strengths and weaknesses relative to life demands, and is the most pragmatic and objective of the assessment process" (McCue, 1994). Unlike NP tests functional assessments are a measure of what a person is capable of or incapable of doing in specific situations, under certain conditions and given unique demands. The HTT attempts to capture this real-world, multi-tasking component by requiring the participant to recall information, actively move objects, and maintain focus

throughout the assessment. The HTT assessment tool used in this study altered clinical testing constraints by allowing the subject to: (a) potentially create more structure when initiated, (b) be unconstrained by a time limit, (c) self evaluate their performance on the task, and (d) go back and self correct without feedback from the examiner.

Scoring of the HTT is specific to each individual step, which helps to pinpoint where in the HTT process the individual loses his/her ability to organize, plan and self regulate. The scoring allows for the qualitative evaluation of the subject's problem solving strategies, abilities and limitations. This approach was used in the design of the scoring form, allowing us to take into account not only how the participants completed the actual HTT, but also how well they initiated and utilized previously learned strategies, followed directions, and exhibited self awareness. EF's are assessed throughout the whole HTT process; this includes watching the HTT task as it is performed and then completing the task from memory using life-like props. After completing the HTT, the participants are provided with a post-assessment which requires them to rate how well they believe they did on the task, recognize if they broke the rules (self-awareness), and state their strategy(s) (planning) and what they would have done differently (evaluation and problem solving).

In recent years, there has been a heightened awareness of the need for tests to produce not only testing scores, but also additional behavioral observations. "In addition to using less structured assessment protocols, executive functions also should be evaluated through observations of the patient in his or her normal activities" (Bennett, 2001, pp. 253-257). We had the opportunity to compare testing scores to the behavioral observations of the participants collected over a 7 week period. Using the Pre-vocational Checklist (Silver et al., 1988) as a

model, the behavioral rating scale provided information concerning each of the participant's current EF abilities. This component of the study allowed for consistent data about each of the subjects, adding an important real-world context that enables examination of real-world behavior across the included domains of EF. Also, it supplied information on how appropriately the HTT relates to current functioning. Bennett (2001) noted the importance of such behavioral observations. "If a person obtains a normal score but in a way that would never be successful in her or his activities of daily living, the person is still impaired psychologically." The location and population of this study allowed for the collection of solid data from a rehabilitation professional concerning an individual's behavior outside of testing or a structured environment such as school and clinical settings. By gathering behavioral data, we were able to provide a more comprehensive understanding of how externally valid the HTT and the standardized tests scores were.

1.4 SPECIFIC AIMS

In an attempt to complement the data from NP testing with an ecologically valid functional assessment of EF, four specific aims have been incorporated into this pilot study's design: (a) to standardized administration of the HTT and determine preliminary scoring reliability of this functional assessment, (b) to determine if the scores from HTT are related to scores from standardized NP tests know to measure similar EF's, (c) to identify where most errors and self corrections are made to help standardize the administration and scoring of the HTT, and (d) to determine the relationship between performance of the HTT and measures of real world EF that the individuals exhibit on a day to day basis in varying social and environmental settings.

2.0 METHODOLOGY

2.1 RATIONALE FOR RESEARCH DESIGN

The neuropsychological tests were chosen due to their extensive use in the measurement of executive functioning (EF) as well as their use in studies that look at ecological validity of these and other tests (Burgess et al, 1998). The hamburger turning task was selected as the simulation assessment based on its real-world origin and prior use in CSEP (Cognitive Skills Enhancement Program) as a rehabilitation tool. The behavioral rating scale is a modified version of the N.Y.U. Pre-vocational Checklist (Silver, et al, 1988), which has been used in related studies and in rehabilitation settings. The questionnaire was modified to include only items relevant to this study.

2.1.1 Data Management

Raw data was kept in on a password protected computer; paper copies of the measures and NP tests were labeled using a random number and letter of the alphabet. A double entry was performed on a subset of the subject's scores (t-scores and raw) to verify that the data entry and conversion was entered correctly. Raw data on the NP tests were converted to standard scores using an accepted normative conversions.

2.1.2 Participants

A sample of 23 cognitively disabled individuals, ages 18 to 26 years, participated in this study. The subjects included 17 males (*M* age = 21, SD= 2.5) and 6 females (*M age* = 20, SD= 1.3) with a range of cognitive disabilities, including pervasive developmental disorder, acquired brain injury, autism spectrum, learning disability, attention deficit disorder, and other neurological disorders. Based on the information from the subjects medical records the mean full scale IQ (FSIQ) of the participants was 84.8. (SD = 9.2). For this convenience sample, we recruited subjects from the CSEP program, which is staffed by the University of Pittsburgh and located at the Hiram G. Andrews Center in Johnstown, PA. The inclusion criteria were as follows: subjects had been admitted to HGA and the CSEP program, were between the ages of 18 and 35, had no physical or sensory impairments that might significantly impede HTT performance, did not require proxy guardianship, and were able to comprehend the implications of participating in this study. The exclusion criteria were consistent with the CSEP program screening requirements, which are as follows: (a) a full scale IQ of 70 or greater, (b) no active drug and alcohol abuse, (c) no active psychosis or other mental health issues that would preclude completion of the study, and (d) no major behavioral disturbances that would preclude completion of the study. The CSEP screening process excludes all individuals who are decisionally impaired, and thus, they were not considered for this study.

Table 1. Primary Diagnosis of Subjects

Subjects		
%	N	Primary Disability Categorization
35%	8	ADHD
17%	4	Traumatic Brain Injury
44%	10	Learning Disability
4%	1	Autism Spectrum

2.1.3 Assessment conditions

Participants were randomly assigned to one of two conditions to control for practice and cueing effects between the two groups, those who were administered the HTT prior to the NP tests and vice versa (sees Figure 1). Condition A required the participants to complete all of the psychometric testing prior to the HTT. Condition B required the participants to complete the HTT first and then complete the psychometric tests.

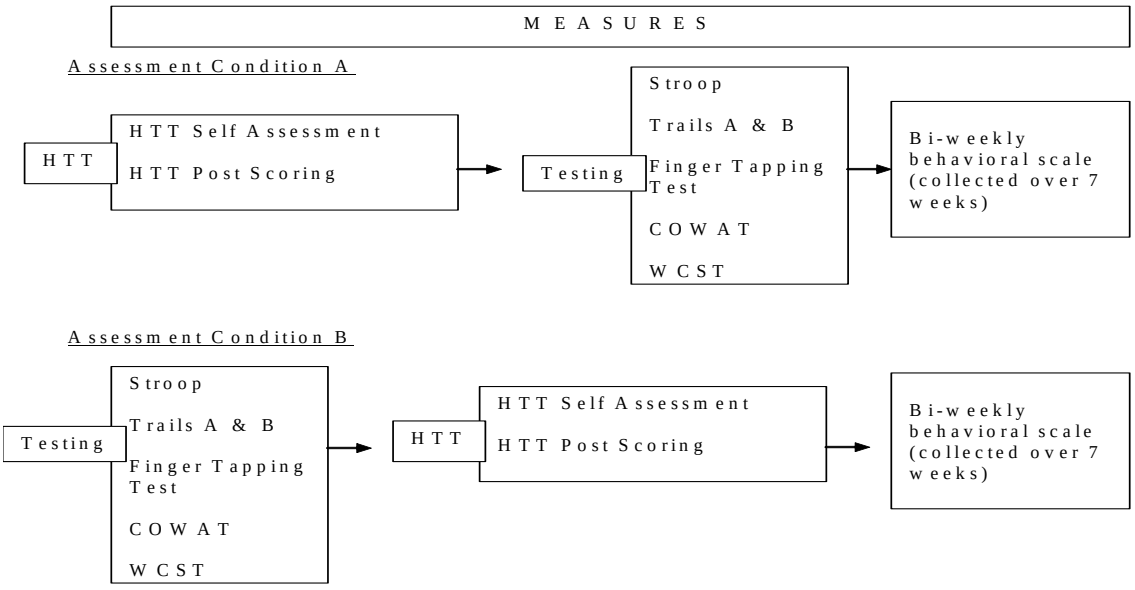

Figure 1 Assessment Conditions

2.1.4 Measurements

The following measurements were used in this pilot study. The measurements are grouped into the following sections, neuropsychological tests, HTT measures (used during the administration) and the behavioral measures.

2.1.4.1 Neuropsychological tests

The Wisconsin Card Sorting Task (computerized version) (Milner, 1963; Nelson, 1976; Struss et al., 2000): 120 cards are presented with 4 symbols on each, and the individual is asked to match the cards according to how they think they should be matched. This task is thought to measure abstract behavior and flexibility, by requiring subjects to shift multiple times during the task.

F-A-S (Crawford, Moore & Cameron, 1992; Spreen & Strauss, 1998), F-A-S: Three trials, each trial requires the participant to recall words that begin with one of the letters F, A, and S for 60 seconds per word. The overall score for all three trials is the sum of all the permissible words produced. The scores of this test reflect flexibility in processing and verbal fluency related to the left frontal lobe (Ruff & Schraa, 2001).

Finger Tapping (Halstead, 1947; Reitan and Wolfson, 1993; Spreen and Strauss, 1998): Examines fine finger dexterity, motor speed, and right/left body comparisons. The subject uses a mounted tapper equipped with a counter. There are several trials administered until the criterion was met. Studies have shown that both neurological and psychological pathology can significantly impact motor speed, which may limit performance in other functional areas (Wefel, Hoyt & Massama, 1999).

Stroop Color-word test: The Stroop test provides an evaluation of sustained and selective attention and inhibition. The inhibition is measured by requiring the subject to refrain from using an automatic response and instead produce a less automatic one. The test requires the participant to label colors of print which spells out conflicting names of colors. 45 seconds per trial, 3 trials (Stroop, 1935; Jensen & Rohwer, 1966).

Trailmaking Tests A and B: Part A requires visual scanning, numeric sequencing and visuomotor speed (Reitan, 1958). 25 circles are randomly dispersed over a sheet of paper and numbered from 1 to 25. Part B is set up the same way but integrates the numbers 1-13 and letters A-L. The subject is to sequentially connect these circles following specific rules. The scoring procedure was the same as was used by Reitan (1958), which equals the total time taken for the participant to complete each part.

2.1.4.2 HTT measures

The hamburger turning task (HTT): The HTT requires the participant to watch a video of an individual verbally explaining each step while performing the HTT (see Appendix A). They are instructed to use any strategy they wish in order to remember this information while watching the video. Once the video concludes, the participant is then introduced to the HTT props (faux grill, spatula, salt, faux hamburgers). In addition to being familiarized to the props, they are also informed that they are to treat the grill and other props as if they are real. For example, if an individual rested their hand on the surface of the grill this would be considered a rule break, and scored as such. It is also explained that they are to concentrate on accuracy rather than speed. The administrator videotapes this process in order to allow for accurate post-scoring. When the participant has completed the HTT, they are then asked to fill out a self-assessment. The administrator then watches the recorded video of the participant completing the HTT and scores him/her on the post-HTT scoring sheet. Instructions on how to use the scoring sheet and how to score specific errors (ex. adding a row) can be seen in Appendix D.

HTT post-scoring sheet: The HTT-post scoring sheet breaks the HTT down into specific steps (see Appendix B). This 15 page form scores a correct step with a 2, skipped steps -1, rule breaks -1, added steps-1, and self-corrected improper sequence (but correct action) is 1. The scores created by this measure are a raw score; a perfect score is 288. To determine the interrater reliability three individuals were provided instruction on how to score the HTT (see Appendix E), and then independently watched five randomly selected subjects performing the HTT on video.

HTT self-assessment form: Consists of four questions that the participant is asked to fill out at the conclusion of the HTT (see Appendix C). Each answer was scored as accurate 2,

moderately accurate 1, or completely wrong 0. A score of all 2's would reflect a person who has exhibited accurate self-awareness, is able to evaluate their performance accurately, and has created a strategy for completing the HTT. A score of all 0's would reflect a person who was completely inaccurate in their self-assessment, was unable or unwilling to complete the answer, and did not generate a strategy for the task. Those whose scores were primarily 1's would fall in-between complete awareness and those who were unaware. For example; individual A only made three mistakes on the HTT and rated himself as "good"; he would have been scored as a 2. Individual B made multiple mistakes and broke three rules during the HTT and rated himself as "excellent"; he was scored as a 0. Individual C did not make any mistakes on the HTT but only scored himself as "good"; he was not completely accurate, but he was also not completely wrong.

2.1.4.3 Behavioral Measures

The Behavioral rating scale: In addition to these testing situations, the participants' daily behaviors were recorded by one of the co-investigators, a cognitive therapist, or a rehabilitation specialist, on a bi-weekly basis over a 7 week period using the behavioral rating scale (see appendix D). The behavioral scale includes four sections related to executive functioning: 1) ability to initiate and persistence, 2) attention and concentration, 3) memory functions, and 4) executive functioning. Each of these categories is not discrete. Collection of this data provided us with consistent insight about each of the participant's behavior over a 7 week period of time in differing environments. The N.Y.U. Pre-vocational Checklist was used as a model for the structure and content (Silver, et al, 1988).

Table 2. Measures

Measures	Cognitive Demands (McCue, 1994)
WCST	Perseveration, abstract thinking, cognitive flexibility, problem solving, using feedback on performance and concept formation
FAS	Verbal fluency and spontaneity
FTT	Motor speed and motor control
Stroop	Sustaining selective attention, speed of processing, reading speed and inhibition.
Trailmaking	Speed, attention, sequencing, shifting alternating sets and mental flexibility
HTT Post Scoring	Create a single score for the HTT
HTT Self-Assessment	Provide insight into the participant's ability to implement strategies, self-evaluate and problem solve.
Behavioral Ratings Scale	Measures each participant's ability to maintain initiation and persistence, attention and concentration, memory and executive function over a 7 week period of time. Providing a comprehensive evaluation of each participant's level of functioning and behaviors on these levels.

2.1.5 Analysis

In order to accomplish the specifics aims of this study the following analyses were run on the data collected over the course of the study. The specific aims included: 1) To standardized administration of the HTT and determine preliminary scoring reliability of this functional assessment To establish reliability of scoring on the HTT and its ability to be used as an appropriate measure, an interrater reliability correlation was run on four independent raters. Comparing ratings across 4 independent raters; all raters scored five HTT administration videos independent of one another. 2) To determine if the scores from HTT are related to scores from standardized NP tests know to measure similar EF's, In an attempt to establish preliminary validity and the significance of the HTT as an assessment tool for executive functioning, scores

from standardized neuropsychological tests and the HTT were compared using a Pearson Product correlation. 3) To identify where most errors and self corrections are made to help standardize the administration and scoring of the HTT a frequency was generated and the percentages of each of the score types was calculated. 4) To determine the relationship between performance of the HTT and measures of real world EF that the individuals exhibit on a day to day basis in varying social and environmental settings. To address ecological validity, or how the HTT relates to real-world functioning, the scores on the HTT were compared to behavioral data collected over 7 weeks. A Pearson Product correlation was also run between the neuropsychological tests and the behavioral data to determine if there was a significant relationship. T-scores were generated for all of the neuropsychological tests, and raw scores were generated for the three hamburger turning task outcomes. There were no significant differences found between the two assessment conditions, those who were administered the HTT first and then the NP tests, and vice versa.

3.0 RESULTS

The results are presented in four areas: 1) preliminary interrater reliability, 2) the relationship of the HTT and scores of the neuropsychological tests; 3) the relationship of the HTT and scores on the neuropsychological tests to the behavioral data; 4) trends and relationships in the data.

3.1 STATISTICAL ANALYSIS

3.1.1 Preliminary interrater reliability of HTT-post scoring form.

The preliminary interrater reliability was examined by comparing the scoring of the HTT performance of 5 subjects, by four trained individuals who rated performances independently of one another. A Pearson correlation-coefficient was run on this normally distributed data. The preliminary reliability coefficients for scoring the HTT range from .994 to .998, all had a p-value of .000 (see Table 3). An Interclass Correlation-coefficient generated a Cronbach's alpha of .999.

Table 3. Interrater Reliability

		Rater 1	Rater 2	Rater 3	Rater 4
Rater 1	Pearson Correlation	1	.998(**)	.996(**)	.998(**)
	Sig. (2-tailed)		.000	.000	.000
	N	5	5	5	5
Rater 2	Pearson Correlation	.998(**)	1	.997(**)	.994(**)
	Sig. (2-tailed)	.000		.000	.000
	N	5	5	5	5
Rater 3	Pearson Correlation	.996(**)	.997(**)	1	.995(**)
	Sig. (2-tailed)	.000	.000		.000
	N	5	5	5	5
Rater 4	Pearson Correlation	.998(**)	.994(**)	.995(**)	1
	Sig. (2-tailed)	.000	.000	.000	
	N	5	5	5	5

3.1.2 HTT Results

The HTT scores ranged from 24 to 288 (no errors made) with a mean score of 165.26 (SD = 66.13). The total number of co-mission (adding steps), omission (forgetting a step), rule breaks and sequencing errors are displayed in Table 4. The frequencies of the scores on the HTT are shown in Table 5. A one-way ANOVA and LSD post hoc analyses were run on the primary diagnosis and the HTT scores of each of the subjects. One subject had a diagnosis of Autism Spectrum Disorder this was dropped from the analysis and coded as a missing variable. No significant relationships were found between the primary disability classification and HTT performance.

Table 4. HTT Errors

Type of Error	Percentage
Co-mission	30%
Omission	52%
Sequencing	16%
Rule Breaks	2%

Table 5. Frequency Distribution of HTT Scores

HTT Scores	Frequency	N	Percent
0-50	1	1	4.3%
50-100	3	3	13.0%
100-150	4	4	17.4%
150-200	10	10	43.5%
200-250	2	2	8.7%
250-288	3	3	13.0%

3.1.3 HTT vs. Neuropsychological tests

The degree of association between the hamburger turning task (HTT) and the neuropsychological tests is presented in Table 6 ($p<.05$). All of the data except the Finger Tapping Test for the dominant hand was normally distributed. A Pearson product-moment correlation coefficient was run on scores which were normally distributed. There were relationships between the HTT and the Stroop Color Word (.614**), Wisconsin Card Sorting Task Perseverative Response (.469*), the Wisconsin Card Sorting Task Perseverative Errors (.447*). This indicates that there is a relationship between a higher score (reflecting better performance) on these neuropsychological

tests of executive functions that look at attention, flexibility, planning, and inhibition was related to a higher score on the HTT (reflecting better performance).

Table 6. Correlation of HTT and NP Scores

	WSCT perseverative errors	WSCT perseverative response	Stroop Color Word
HTT Post	.447*	.469*	.614**
P values	.033	.024	.002

3.1.4 HTT vs. behavioral rating scale

A Pearson correlation coefficient was used to compare the HTT performance and behavioral data, both of which were normally distributed. Negative correlations denote good concordance between the measures, as a high HTT score indicates good performance, whereas a high behavioral data score represents poor performance. The Hamburger Turning Task and the behavioral data are significantly correlated (-.585**, P value=.003). This indicates a strong relationship between the HTT and the behavioral measure scores.

3.1.5 Neuropsychological tests vs. behavioral rating scale

The only relationship between the behavioral data and the NP tests was with the Stroop Color Word (-.552**) and the Stroop word (-.423*). A high score on the Stroop Color Word indicates high functioning, and a high score on the behavioral rating scale represents low functioning.

Table 7. NP tests and Behavioral Data

	Behavioral Rating score	Stroop color word	Stroop word
Pearson Correlation	1	-.552(**)	-.423(*)
Sig. (2-tailed)		.006	.044
N	23	23	23

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

3.1.6 Trends in the data

The means for performance of the HTT and neuropsychological testing for the 23 subjects are presented in Table 8.

Table 8. Means and SD of Testing and Demographic Data

Variables	**Mean**	**SD**
Age	20.65	2.208
HTT score	165.26	66.137
Trails A Time	36.30	9.266
Trails B Time	39.87	12.686
FAS Total	36.30	9.266
Stroop Word	37.48	7.242
Stroop Color	38.57	9.28
Stroop Color Word	39.87	9.27
FTT dominant	35.91	7.66
FTT non-dominant	36.77	12.88
Wisconsin Card Sorting Task total errors	49.78	14.49
Wisconsin Card Sorting Task perseverative errors	52.96	18.11
Wisconsin Card Sorting task perseverative responses	53.48	18.89

There is a relationship between the subject's full scale IQ and their performance on the HTT, accuracy on the self assessment, and functioning on the behavioral rating scale (see Table 9). A correlation between the self-assessment and all of the other variables is located in Table 10. In general, performance on the HTT, behavioral data and the FSIQ were associated with self-assessment in that poorer performance on measures the less accurate the sel-assessment.

Table 9. HTT measures and FSIQ

Test or Assessment	FSIQ	p-values
HTT Post	.657**	.001
Self Assessment	.500*	.013
Average of Behavioral data	-.668**	.003

** Correlation is significant at the 0.01 level (2-tailed).
* Correlation is significant at the 0.05 level (2-tailed).

Table 10. Correlations to Self Assessment

Correlation to Self Assessment Measure			
Variable	Correlation to Self Assessment	P value	
HTT Post	.516*	.012	Pearson Product Correlation
FSIQ	.500*	.015	
Behavioral Data	-.722**	.000	
FAS	.182	.405	
Stroop Color Word	.430*	.040	
Stroop Color	.285	.188	
WCST Perseverative Errors	.246	.258	
WCST Perseverative Errors	.266	.220	
WCST Errors	.245	.260	
Trailmaking A	.189	.389	
Trailmaking B	.243	.264	

Finger Tapping D	.245	.261	Spearman rho
Finger Tapping ND	.132	.559	

There were no significant differences found between the two assessment conditions, those who were administered the HTT first and then the NP tests, and vice versa.

4.0 DISCUSSION

4.1 INFERENCES FROM THE DATA ANALYSIS

The present investigation sought to establish a consistent scoring protocol as well as standardized administration of the HTT. The HTT's ability to assess an individual's executive functioning from day to day, or ecological validity, was explored using behavioral data. The comparison between the HTT and the standardized neuropsychological tests provided information about the components of EF that the HTT measures. The results of this investigation indicate that the HTT may be useful as a functional evaluation for certain aspects of executive functioning in a real-world setting. However, since the HTT was not correlated to three of the five NP tests, we cannot conclude that the HTT necessarily measures the same cognitive process which these tests are standardized. This will need to be examined further in later research.

Across the data, there was a consistent relationship between the HTT, the Stroop Color Word and the behavioral rating scale. The HTT's correlation to the Stroop supports the HTT's validation as a measurement of certain aspects of executive functioning, specifically speed, attention and flexibility. The significant relationship between the HTT and the behavioral data indicates that performance on the HTT is related to how well an individual functions everyday, according to the five levels of executive functioning on the behavioral rating scale (maintain initiation and persistence, attention and concentration, memory and executive functions).

Interestingly, there was only one significant relationship between the NP tests (Stroop Color Word) and the behavioral rating scale, which causes us to question how ecologically valid these other NP tests are, particularly in the areas that the behavioral rating scale measured. This would support previous research that looked at the ecological validity of these NP. Sbordone and Guilmette (1999) found in their review of neuropsychological tests that no single specific test or measure could accurately predict the everyday functioning of brain injured patients. With this information, they stated that neuropsychologists should not rely solely on neuropsychological test scores to predict everyday functioning. This is where functional assessments such as the HTT can be used to supplement NP test data and allow for a more comprehensive understanding of an individual's strengths, weaknesses and overall functioning in all situations.

4.2 RELATIONSHIPS IN ADDITIONAL DATA

There was also a high correlation between an individual's full scale IQ score and the HTT-post scoring sheet, the self assessment and the behavioral rating scale data. There was only one relationship between the self-assessment and any of the NP tests, the Stroop Color Word. These relationships also support the hypothesis that the HTT is an accurate measure of an individual's functioning and self-awareness. We did not specifically address self-awareness in this study, but the self report measure or one of a similar format might prove to be a good indicator of self-awareness when using functional assessments such as the HTT. The subject's diagnosis was not directly related to the score they received on the HTT, which allows us to conclude that this functional assessment tool can be used as a measure for a wide range of abilities and cognitive disorders and does not discriminate between them.

4.3 LIMITATIONS

4.3.1 Administration

One of the significant confounding variables in this study was due to time constraints and logistical issues, the primary investigators was required to administer the HTT, score the HTT, and score each participant on the behavioral rating assessment, creating the potential for an experimental bias to be reflected in the scores. Future studies need to control for this in their design.

The behavioral data gathered for this study was comprehensive (observations of structured, unstructured time by a rehabilitation professional) and will most likely not be replicated in a future study because of the unique and easily observable features of this subject population (due to the onsite living arrangements and constant supervision), but provides this pilot study with a certain level ecological validity. It would have been beneficial to use a standardized and more widely recognized measure for this data, rather than modifying the N.Y.U Pre-vocational Checklist. The Dysexecutive Questionnaire (Wilson, et al., 1996) and the Brock Adaptive Functioning Questionnaire (Dywan & Segalowitz, 1996) are structured rating scales that are used to measure similar behavioral functioning. However, these measures have been designed for individuals with brain injury rather than the broad range of cognitive disabilities present in this study. An interesting addition that future researchers might add to their study design is to have the subject's rate themselves, as well as having an outside party scoring them on the behavioral components. This would supply additional information about self-awareness that we attempted to gain by using the short self-assessment.

Due to the time constraints of this thesis, it was only possible to use the first 7 weeks of behavioral data versus the 16 weeks we had anticipated. The subjects were grouped into two terms, each term lasted 16 weeks. For the first term, it was decided that we would take only the first 7 weeks so the comparison would be equivalent to the second term, rather than using an average of the whole 16 weeks and comparing it to 7 weeks in term 2.

4.3.2 Subjects

The variability of diagnoses and FSIQ of the population and the small sample size of this pilot study do not allow for a great deal of genralizability. Four subjects were asked to leave the rehabilitation program from which this study's participants were recruited, causing a decrease in the already small sample size. Future studies should look at specific diagnoses and compare them to a control group to better determine the use of this functional assessment across populations. The HTT should also be administered and scored for individuals with normal range IQ and no cognitive disabilities to determine what the ranges should be for the scores of the HTT and how many errors are "normal".

4.3.3 Scoring

It was determined that using a scoring sheet while the HTT is being performed was difficult, and steps were often missed while the administrator looked down to record the answers. This may cause logistical and time issues when using the HTT as a functional assessment, because the HTT-post scoring sheet requires a video recording of the subject performing the task, which must be watched after the administration. This was determined as the most accurate way to

score the HTT, because it allows the scorer to replay the video, pause to record answers, and rewind if they miss a step. In this study, we did not look at the time that the subjects took to complete the task. This may be an important addition to the scoring design.

Not all errors are equivalent and that is one issue we did not address in this study. Just as Shallice and Burgess (1991) concluded about the Multiple Errands Test, not all forms of errors are equivalent in terms of their significance. The same is true of the HTT. For example one participant began to flip the burgers with his hands (did not use the spatula), another subject kept dropping burgers off of the grill and onto the floor and there were four subjects that added rows that were not included in the administration video. These errors were scored as having the same weight as a simple error like forgetting a step. There is a need for each of these and other errors to be categorized and scored according to their significance in relation to level of EF.

4.4 FUTURE RESEARCH

In order for the HTT to be used more widely as an assessment tool the following improvements need to be made to the scoring and administration: 1) errors should be scored more precisely in order to determine what specific issues that individual is having; 2) determine a range for the scores (Ex. low functioning a score of = 0-50); and 3) compare scores of those who have a normal FSIQ and no cognitive disabilities to help in recognizing the number and types of errors that are “normal”. In addition to these scoring changes, future research should look at the how well the HTT relates to vocational and rehabilitation outcomes. Another important issue that should be addressed is the ability of the HTT to pick up on visuo-spatial strengths and weaknesses. For instance, there were two subjects who did moderately well on all the NP tests

and the behavioral data but scored as two of the lowest on the HTT due to the errors made while laying out the materials spatially (poor visuo-spatial abilities). This suggests that the HTT may be able to capture deficits in visuo-spatial abilities in addition to the other EF it measures.

4.5 IMPLICATIONS FOR REHABILITATION

These preliminary results suggest that the HTT is able to be reliably scored and has a standardized administration protocol that allows it to be a useful measure of executive functioning of everyday behavior. These results indicate that a standardized administration of the HHT can offer an enhanced, comprehensive picture of an individual's level of current functioning when paired with NP tests. The HTT has the ability to provide insight to a clinician or rehabilitation professional about how an individual might act in a vocational setting and allow the professional to pick up on behaviors that may not be seen in the "paper pencil" format of neuropsychological tests.

In the past, the HTT has been used in a rehabilitation program as a way to teach self-awareness and implement strategies for procedural memory. Since we created a format for administration and a scoring protocol, the HTT can be used initially as a standardized functional assessment and later be developed as a tool to assist rehabilitation professionals in teaching subjects how to learn processes, memory strategies, and self awareness. Rehabilitation professionals need to be equipped with more resources and tools similar to the HTT to create a better picture of real-world functioning on a variety of hands on tasks. This is important to the individuals they are serving, as well as their caretakers and case managers in helping their clients set reasonable goals and determine what strengths and weaknesses that individuals possesses. A

number of the subjects have begun to recognize their own strengths and weakness as a direct result of their performance on the HTT. The format of this functional assessment allows the subject to see specifically where they failed, what mistakes were made and what caused those mistakes (memory, attention). Seeing this in a real world task has the potential to be much more profound than scoring poorly on a test, missing an appointment, or not remembering information from class.

A uniformly accepted definition of executive functions has not been established and assessments of these complex cognitive processes are not always comprehensive or ecologically valid, but the impact that they have on an individual's ability to function is indisputable. One of the specific aims of this study was to determine what aspects of EF are measured by HTT . These include: 1) the ability to sequence, 2) following rules, 3) forgetting steps (omission), 4) adding steps (co-mission), 5) self-evaluation and use of strategies (on the post self-assessment), 6) memory and 7) procedure. In addition to these areas of executive functioning (EF) the HTT has the ability to provide a comprehensive picture of an individual's functioning in a number of different areas. 1) How well can they follow through on a task they are unfamiliar with? Are they able to complete the task when they are not cued as to how well they are doing and cannot ask for assistance? 2) How well can they correct mistakes, and are they able to do so in an appropriate manner. For example, one subject dropped a burger on the floor and then picked it up and placed it back on the grill, had this been a real life situation they would have contaminating the food and would most likely be reprimanded by their superior. 3) Are they able to create a strategy? Can they recognize when they will require a strategy in order to complete a job (ex. way to memorize steps), and are they able to implement this strategy and follow through with it on the task? 4) How do they react when they are conscious that a mistake

has been made? There were two subjects in this study who became frustrated and quit the task early when they became aware that the task was not being performed correctly. Behaviors such as these are important to observe and provide insight about that individual's ability to deal with stressful situations. This additional information about EF that can be gleaned from simulation tasks like the HTT is important for setting and planning for rehabilitation professionals to address, and are pertinent to rehabilitation, educational and vocational goals. Future research should attempt to determine a more objective ways to measure these critical aspects of EF rather than looking at one or two individual's functions.

In summary, this study's primary intention was to provide another practical option for clinicians and rehabilitation professionals to use in helping to determine an individual's executive functioning outside of the office. HTT has been found to be an ecologically valid measure in the following areas: 1) ability to initiate and persistence, 2) attention and concentration, 3) memory functions, and 4) executive functioning. Many individuals with deficits in specific areas of EF are passing through NP assessments, and even vocational and functional assessments, without clinicians having a clear understanding of their deficits. The HTT has the potential to add another resource to assist in finding that point where an individual will fall short in an area of EF, act out in an idiosyncratic way or present a behavior that was not seen in NP or more structured testing.

APPENDIX A

VERBAL STEPS OF THE HTT ADMINISTRATION

1. Place 4 burgers on grill
2. Salt the first row
3. Flip the 1st row to 2nd row
4. Place next set of 4 burgers onto row 1
5. Salt all burgers in rows 1 and 2
6. Press burgers in row 2
7. Flip 2nd row to the 3rd row
8. Flip 1st row to 2nd row
9. Place 4 burgers onto row 1
10. Salt burgers in row 1 and 2
11. Press burgers in row 2 and 3
12. Flip 3rd row to 4th
13. Flip 2nd row to 3rd
14. Flip 1st row to 2nd
15. Place last 4 burgers onto row 1.
16. Salt burgers in row 1 and 2
17. Press burgers in row 2 and 3
18. Stack burgers in row 4
19. Flip burgers in 3rd row to 4th row
20. Flip burgers in 2nd row to 3rd row.
21. Flip burgers in 1st row to 2nd .
22. Salt the burgers in row 2
23. Press the burgers in row 2 and 3.
24. Stack 4th row.
25. Flip the 3rd row to the 4th row
26. Flip the 2nd row to the 3rd.
27. Press the burgers in row 3
28. Stack the 4th row
29. Flip the burgers from the 3rd row to the 4th
30. Stack the 4th row.

Figure 2 Steps of HTT

APPENDIX B

Post HTT Scoring Sheet using video: Page 1 of 15

Row 1

1st Burger Set

	Burger 1 row 1	Burger 2 row 1	Burger 3 row 1	Burger 4 row 1
Place on grill (2)				
Salt burger (2)				
Flip to row 2 (2)				
Salt out of sequence, but still in same row (+1)				
Salt 2+ times (-1)				
Salting other: Explain				
Forget to salt (-1)				
Flip out of sequence but still correct (+1)				
Flip 2+ times (-1)				
Forget to flip (-1)				
Flipping other: Explain				
Rule Break (-1) Explain:				
Total				

Figure 3 HTT Post-Scoring Sheet

APPENDIX C

POST HTT SELF EVALUATION

On a scale of 1 to 5 how well do you think you did on this Hamburger Turning Task?

1	2	3	4	5
Failed	poor	ok	good	perfect

Did you break any of the rules of the HTT? (explain)

What was your strategy for completing this task?

What would you have done differently if you were to do this task again?

APPENDIX D

BEHAVIORAL RATING SCALE

Behavioral rating scale

Name:_____ Date of Evaluation_____

Examiner:_____

1=complete ability, independent

2= Adequate ability, minimal intervention

3= Moderate ability

4= Minimal ability, maximum structure

5=total inability

Ability to Initiate and Persistence

_____1. The individual's initiation to the task.

_____2. Ability to recognize their own limitations.

_____3. Maintained task-relevant behavior

_____4. Ability to disengage appropriately from performance of task

_____ Total

Attention and Concentration

_____1. Is able to regulate impulsivity

_____2. Ability to survey and orient to the task

_____3. Level of internal distractions

_____4. Level of external distraction

_____ Total

Memory Functions

_____1. Recall of routine sequential task components during performance of the task.

_____2. Ability to remember the implicit and explicit rules of the task.

_____3. Ability to reconstruct from memory task routine

_____4. Recall and adherence to all of the rules for HTT

_____ Total

Executive Functioning Skills

Goal Identification/strategy.

_____1. Understanding formulation of a verbal/explicit goal

_____2. Understand formulation of an implicit or inferred goal

_____3. Able to generate a strategy and plan of approach

Executive Skills/Planning/Self-monitoring

_____1. Ability to detect errors

_____2. Ability to correct errors

_____3. Ability to make inferences from specific errors corrected

_____4. Ability to set priorities

_____5. Ability to organize steps and materials

_____6. Ability to attend to own task performance

Problem Solving Skills

_____7. Ability to sequence

_____8. Ability to master routines

_____9. Ability to generalize/transfer information to a new situation

_____10. Ability to follow rules

____ Total

_____ Overall Total

APPENDIX E

INSTRUCTIONS ON THE HTT POST-SCORING SHEET

1. All correct actions are a 2. Score each burger individually.
2. All incorrect actions are scored as a -1. This includes forgetting a step and adding a step.
3. Rule breaks include anything that would potentially hurt the individual performing the task if the grill was real (ex. Placing hand on grills surface), or would be harmful to the burgers (ex. Leaving the paper on the burgers). Additional rule breaks: setting spatula on grill, leaning against the grill, allowing a burger to roll off the grill, picking up the burgers with hands, and flipping burgers with hands.
4. When a row is stacked too early score this as “forgot to flip”(-1) on the page that they should have flipped and instead stacked and also score them as “skipped step” (-1) on the page that scores them for stacking.
5. When rows are added score them as “forgot to stack” (-1) and “adding step” (-1) for each burger. Do not give them a point for stacking out of sequence.
6. If a subject pats out of sequence (ex prior to salting) only mark the patting out of sequence. So only mark the first step that is out of sequence wrong.

** When training other how to use the scoring sheet we watched a practice video where they could ask questions and learn how to watch the task appropriately.

BIBLIOGRAPHY

Anderson, S.W., & Tranel, D. (1989). Awareness of disease states followinf cerebral infarction, dementia, and head trauma: Standardized assessment. *The Clinical Neuropsychologist, 3,* 327-339.

Bennett, T.L. (2001). Neuropsychological evaluation in Rehabilitation planning and evaluation of functional skills. *Archives of Clinical Neuropsychology, 16,* 237-253.

Bryan, J. & Luszcz, M. (2000). Measurement of executive function: Considerations for detecting adult age differences. *Journal of Clinical and Experimental Neuropsychology 22(1),* 40-55.

Benton, A.L. & Hamsher, K. deS. (1989). *Multilingual Aphasia Examination.* Iowa City: AJA Associates.

Burgess, P.W., et al. (1998). The ecological validity of tests of executive function. *Journal of International Neuropsychological Society, 4, 547-558.*

Butler, R., Anderson, L. Furst, C.J., Namerow, N. & Satz, P. (1989). Behavioral a ssessment in neuropsychological rehabilitation: A method for measuring vocational-related skills. *The Clinical Neuropsychologist, 3,* 235-243.

Cicerone, K.D. & Giacino, J.T. (1992). Remediation of executive functioning deficits after traumatic brain injury. *Neurorehabilitation. 2(3),* 12-22.

Crawford, J.R., Moore, J.W. & Cameron, I.M. (1992). Verbal fluency: A NART-based Equation for the estimation of pre-morbid performance. *British Journal of Clinical Psychology, 31,* 327-329.

Cripe, L.I. (1996). The ecological validity of executive function testing. In Sbordone, R.J., and Long, C.J. (Eds.), *Ecological Validity of Neuropsychological Testing* (pp. 171-202). Delray Beach, FL: GR Press/St. Lucia Press.

Dywan, J. & Segalowitz, S. (1996). Self and family ratings of adaptive behavior after traumatic brain injury: Psychometric scores and frontally generated ERP's. *Journal of Head Trauma Rehabilitation. 11, 79-95.*

Franzen, M.D., & Wilhelm, K.L. (1996). Conceptual foundations of ecological validity in neuropsychology. In Sbordone, R.J., and Long, C.J. (Eds.), *Ecological Validity of Neuropsychological Testing* (pp. 91-112). Delray Beach, FL: GR Press/St. Lucia Press.

Halstead, W.C. (1947). *Brain and Intelligence.* Chicago: University of Chicago Press.

Heaton, R. & Crowley, T. (1981). Effects of psychiatric disorders and their somatic treatments on neuropsychological test results. In S. Filskov & T. Boll (Eds.), *Handbook of Clinical Neuropsychology* (pp. 481-525). New York: Wiley.

Jackson, H.F., & Hopewell, C.A. et al. (1992). The katz adjustment scale: Modification for use of victims with of traumatic brain and spinal injury. *Brain Injury. 6,* 109-127.

Jensen, A.R. & Rower, W.D. (1966). The Stroop Color-Word Test: a review. *Acta Psychologica,* 25, 36-93.

Lezak, M.D., Howieson, D.B., & Loring, D. et al., Neuropsychological Assessment (4th ed.) [Book, Authored Book] (2004). x*iv, 1016pp.* New York: Oxford University Press.

Lezak, M.D. (1983). *Neuropsychological assessment* (2nd Ed.). New York: Oxford University Press, New York.

Lezak, M.D. (1982) The problem with assessing executive functions. *International Journal of Psychology. Vol 17(2-3),* 281-297.

Luria, A.L. (1966). *Higher cortical functions in man.* New York: Basic Books.

McCue, M. et al. (1994). *Functional assessment of individuals with cognitive disabilities: A desk reference for rehabilitation.* Pittsburgh, PA: Center for Applied Neuropsychology.

Milner, B. (1963). Effects of different brain lesions on card sorting. *Archives of Neurology, 9,* 90-100.

Nelson, H.E. (1976). A modified card sorting test sensitive to frontal lobe defects. *Cortex, 12,* 313-324.

Rabbit, P. (1997). Introduction: Methodologies and models in the study of executive function. In P. Rabbit (Ed.), *Methodology of frontal and executive function* (pp.1-38). UK: Psychology Press.

Reitan, R.M. & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical applications.* (2nd ed.). Tucson, AZ: Neuropsychological Press.

Reitan, R.M.. (1958). Validity of the trail making test as an indicator of organic brain damage. *Perceptual and Motor Skills, 81, 271-276.*

Ruff, R.M. & Schraa, J.C. (2001). Neuropsychological assessment. In B. Brian (Ed.), *Handbook of Measurement and Evaluation in Rehabilitation* (3rd Ed.) (pp.233-254). Austin, TX: Pro-Ed.

Sbordone, R. J. & Guilmette, T.J. (1999). Ecological validity: Prediction of everyday and vocational functioning from neuropsychological test data. In JJ. Sweet (Ed.), *Forensic Neuropsychology: Fundamentals and Practice* (pp. 227-254). Netherlands: Swets & Zeitling.

Sbordone, R.J. (2001). Limitations of neuropsychological testing to predict the cognitive and behavioral functioning of persons with brain injury in real-world settings. *Neurorehabilitation 16(4),* 199-201.

Sbordone, R. & Long, C. (Eds.) (1996) *Ecological Validity of Neuropsychological Testing.* Delray Beach, FL: St. Lucia Press.

Shallice. T. (1982). Specific impairments of planning. *Proceedings of the Royal Society of London, Series B, 298,* 199-209.

Shallice, T. & Burgess, P.W. (1991). Deficits in strategy application following frontal Lobe damage in man. *Brain, 114,* 727-741.

Silver, S.M., Ezrachi, O., Kay, T., Rattock, J., Piasetsky, E., & Ben-Yishay, Y. (1988). Administration manual for the N.Y.U. Prevocational Checklist. New York University Medical Center.

Spreen, O. & Strauss, E. (1998). *A compendium of neuropsychological tests* (2nd ed). New York: Oxford University Press.

Stroop, J.R. (1935). Studies of interference in serial verbal reactions. *Journal of Experimental Psychology, 18,* 643-662.

Stuss, D.T., Binns, M.A., Carruth, F.G., et al. (2000). Prediction of recovery of continuous memory after traumatic brain injury. *Neurology, 54,* 1337-1344.

Wong, B.Y.L. (1985). Metacognitive and learning disabilities. In D.L. Forrest-Pressley, G.E. MacKinnon & T.G. Waller (Eds.), *Metacognition, cognition and human performance: Vol. 2. Instructional practices* (pp.137-180). New York: Plenum Press.

Wefel, J.S., Hoyt, B.D., & Massama, P.J. (1999). Neuropsychological functioning in depressed versus non-depressed participants with Alzheimer's disease. *Clinical Neuropsychologist*, 13, (3), 249-257

Wilson, B.A., Alderman, N., Burgess, P., Emslie, H. & Evans, J. (1996). *Behavioral Assessment of Dysexecutive Syndrome.* St. Edmunds, UK: Thames Valley Test Company Limited.