Published in final edited form as:

Neuropsychol Rehabil. 2017 July ; 27(5): 707–721. doi:10.1080/09602011.2015.1037771.

The Reliability and Validity of the Complex Task Performance Assessment: A Performance-Based Assessment of Executive Function

Timothy J. Wolf, OTD, MSCI, OTR/L^{1,2}, Abigail Dahl, BS, MSOT, Colleen Auen, OTD, OTR/L, and Meghan Doherty, MSOT, OTR/L¹

¹ Program in Occupational Therapy, Washington University in St. Louis, School of Medicine

² Department of Neurology, Washington University in St. Louis, School of Medicine

Abstract

The objective of this study was to evaluate the inter-rater reliability, test-retest reliability, concurrent validity, and discriminant validity of the Complex Task Performance Assessment (CTPA): An ecologically-valid performance-based assessment of executive function. Community control participants ($n = 20$) and individuals with mild stroke ($n = 14$) participated in this study. All participants completed the CTPA and a battery of cognitive assessments at initial testing. The control participants completed the CTPA at two different times one week apart. The intra-class correlation coefficient (ICC) for inter-rater reliability for the total score on the CTPA was 0.991. The ICCs for all of the sub scores of the CTPA were also high (0.889-0.977). The CTPA total score was significantly correlated to Condition 4 of the DKEFS Color-Word Interference Test ($p = -0.425$), and the Wechsler Test of Adult Reading ($p = -0.493$). Finally, there were significant differences between control subjects and individuals with mild stroke on the total score of the CTPA ($p = 0.007$) and all sub scores except interpretation failures and total items incorrect. These results are also consistent with other current executive function performance-based assessments and indicate that the CTPA is a reliable and valid performance-based measure of executive function.

Keywords

Cognition; Executive Function; Rehabilitation; Ecological Validity; Occupational Therapy

Introduction

In recent years there have been several studies that have reported that the nature of stroke is changing in terms of the age when people experience stroke and the severity of the neurological symptoms. In a study of 7,740 people after stroke, 45% were age 65 or over while 27% were age 55-64 (T. J. Wolf, Baum, & Connor, 2009). These findings support

Corresponding Author: Timothy J. Wolf, Program in Occupational Therapy, Department of Neurology, Washington University School of Medicine, 4444 Forest Park, St. Louis, MO 63108, Phone 314-286-1683, Fax 314-286-1601, wolft@wustl.edu.

The authors have no financial interest in the results of this study. The assessment evaluated, the Complex Task Performance Assessment, is available in the public domain at ot.wustl.edu.

other research that suggests that those affected by stroke are younger than previously thought (Hartman-Maeir, Soroker, Ring, Avni, & Katz, 2007; Medin, Barajas, & Ekberg, 2006). These studies are also consistent with a recent Centers for Disease Control and Prevention study that found that ischemic stroke rates rose as much as 50% between 1996 and 2008 in young adults (George, Tong, Kuklina, & Labarthe, 2011). In addition, the severity of neurological impairment post-stroke has decreased. In the same study by Wolf and colleagues (2009), 49% of the population sustained a mild stroke with a National Institute of Health Stroke Scale (NIHSS) score of less than six. The NIHSS is a 15-item scale used to assess the neurological impairment of stroke (Brott et al., 1989). Individuals with a mild stroke (NIHSS <6) are typically independent with ADL at time of discharge from the acute hospital and often don't have many if any overt neurological symptoms. Since these individuals are mostly independent with ADL, the majority are discharged home with little to no rehabilitation services (Rochette, Desrosiers, Bravo, St-Cyr-Tribble, & Bourget, 2007; T. J. Wolf et al., 2009).

Although the majority of the mild stroke population is being discharged home with limited services, research suggests that this population has difficulty reintegrating into their previous activities when they return home and attempt to return to work (Edwards, Hahn, Baum, & Dromerick, 2006); (Hartman-Maeir et al., 2007; Hildebrand, Brewer, & Wolf, 2012; O'Brien & Wolf, 2010) Several investigations have discovered that executive dysfunction has been linked to a decrease in return to work status of individuals post mild stroke (Alaszewski, Alaszewski, Potter, & Penhale, 2007; Medin et al., 2006; O'Brien & Wolf, 2010; Wolfenden & Grace, 2009) . Executive function is required for complex, novel, everyday life activities (i.e., household management, managing finances, work) but not for routine/procedural activities which are typically the focus of acute assessment following stroke (i.e., self-care) (T. J. Wolf et al., 2009). Those with executive function deficits often go undetected because they are able to successfully perform basic activities of daily living but have difficulty with novel tasks and responding to the continually changing demands of home, work, and community engagement (D. F. Edwards et al., 2006). Executive dysfunction is a primary deficit within the mild stroke population and is present within up to 66% of this population (Timothy J. Wolf, Barbee, & White, 2010).

There is increasing need for use of performance-based assessments to detect higher-level cognitive deficits in the mild stroke population. Neuropsychological assessments are considered the gold standard in executive function assessment; however, these assessments have some limitations. They are usually very well structured with clearly delineated goals which contradicts defined components of executive function (Adams, Parsons, Culbertson, & Nixon, 1996). In addition, those with executive dysfunction following mild neurological injury are often able to perform within normal limits on neuropsychological tests, even those tests considered sensitive to executive function (Shallice & Burgess, 1991; Wilson, Alderman, Burgess, Ernsle, & Evans, 2003). Shallice and Burgess (1991) developed the Multiple Errands Test (MET) in an attempt to address limitations with common neuropsychological tests and their ability to identify performance-based deficits in people with executive dysfunction (Shallice & Burgess, 1991). While the MET is prevalent throughout rehabilitation, there are several clinical limitations including limited

psychometric development, limited clinical feasibility (e.g., time and space), and lack of alternate forms to minimize learning effects.

The Complex Task Performance Assessment (CTPA) is a performance-based assessment of executive function that was developed based on the theoretical framework and organizational structure of the MET; the MET is based on the role and action of Supervisory Attentional System in complex multitasking situations (Norman & Shallice, 1986; Shallice & Burgess, 1991). The CTPA, however, was designed specifically to address the limitations of using the MET in a clinical context. The initial pilot study found that the CTPA was able to discriminate between those with neurological injury and community controls; however, the psychometric properties have yet to be evaluated. (T. Wolf, Morrison, & Matheson, 2008). The purpose of this study was to evaluate the psychometric properties of the CTPA, specifically, inter-rater reliability, test-retest reliability, concurrent validity, and discriminant validity.

Methods

Design

The design of this study was: (1) a descriptive, repeated-measures design to evaluate the inter-rater reliability, test-retest reliability, and concurrent validity of the Complex Task Performance Assessment with community controls; and (2) a cross-sectional study with individuals with mild stroke and community controls to evaluate the discriminant validity of the CTPA. The mild stroke participants were recruited from the Cognitive Rehabilitation Research Group stroke registry at Washington University School of Medicine and community controls were recruited using Volunteers for Health, a program sponsored by Washington University School of Medicine to support patient-oriented research. This study was reviewed and approved by the Washington University School of Medicine Human Research Protection Office.

A threshold correlation coefficient of 0.80 was set for this study to establish inter-rater and test-retest reliability. This conservative threshold was established in response to documented practice effects associated with serial administration of cognitive testing. It has been shown in control populations that learning occurs with cognitive testing, specifically on novel tests where a strategy can be identified, and performance improves with multiple administrations (Beglinger et al., 2004).

Personnel training

The CTPA was administered by graduate students in the Program in Occupational Therapy at Washington University in St Louis. Test administrators followed an established training protocol prior to the commencement of this study that included the following: (1) reviewing the testing manual, procedures, and materials with an approved test administrator; (2) administering the assessment on at least five control subjects; and (3) obtaining final approval for test administration by an approved administrator.

Participants

Inter-rater reliability, test-retest reliability and concurrent validity were evaluated using community controls subjects. Inclusion criteria included: (1) adult age (18-80); (2) English speaking; (3) at least high school education; and (4) normal estimated intelligence indicated by a Weschler Test of Adult Reading (WTAR) score greater than or equal to 85. Exclusion criteria included: (1) documented executive dysfunction indicated by a DKEFS scaled score of less than or equal to 7 on Condition 4 of either Color-Word Interference or Trail Making subtest; (2) any history of neurological disease; (3) any history of mental health disorder; (4) any history of cognitive dysfunction; (5) any reported drug/alcohol abuse; (6) profound vision loss (>20/40 after correction); (7) inability to write; and (8) inability to comprehend verbal directions. From the advertisements placed through Volunteers for Health, 40 potential participants responded. Of those 40 respondents, 11 did not meet inclusion/exclusion criteria. Nine individuals were not scheduled to participate due to scheduling conflicts. The remaining 20 participated in the study.

To evaluate discriminant validity a total of 14 individuals with mild stroke were recruited and tested. All participants were admitted to the Barnes-Jewish Hospital stroke service in St. Louis, Missouri. Upon admission, all participants gave consent to be included in the Cognitive Rehabilitation Research Group stroke registry. Participants were asked to give permission to be contacted for future studies related to stroke. Further inclusion criteria were established for this study: (1) a mild stroke (NIHSS scores between 1-5); (2) at least 6 months post-stroke; (3) aged 18-65; (4) documented executive dysfunction indicated by a DKEFS scaled score of less than or equal to 7 on Condition 4 of either Color-Word Interference or Trail Making subtest AND greater than or equal to 7 on Condition 5 of Motor Speed; (5) normal estimated intelligence indicated by a Weschler Test of Adult Reading (WTAR) score greater than or equal to 85; and (6) self-reported ability to write legibly. Exclusion criteria included: (1) history of any other neurological or mental health disorder prior to stroke; (2) history of previous stroke; (3) blindness; and (4) non-English speaking.

Measures

The Complex Task Performance Assessment (CTPA) was the primary outcome measure for this study. This is a neuroperformance-based assessment which simulates the task of working in a library. There are two primary work tasks: (1) Current Inventory Control and (2) Telephone Messaging. These tasks are administered simultaneously.

Current Inventory Control

The Current Inventory Control task requires that the participant calculate the total fine and replacement cost for books and videos that are overdue. The majority of the inventory control is already completed for them; however three titles need to be calculated. The information about each overdue book or video is displayed on a stack of cards. Each card includes the title of the book or video, the name and address of the person who has it checked out, the cost of the fine, and the replacement cost. The participant is told that patrons will pay the lower cost and must determine from the information on the card which cost the patron will pay. In order to complete the inventory control, the participant is asked

to calculate the total replacement costs, total fine costs, and total money to be collected for each title. Participants are instructed to work as quickly and efficiently as possible.

Telephone Messages

The Telephone Message activity requires the person to listen to recorded telephone messages and respond appropriately. There are a total of four messages that are played on a CD or MP3-player. The messages are played randomly throughout the first 15 minutes of the assessment. The messages fall under one of three categories including: declarative, interactive, and reasoning. Declarative messages require the participant to simply record the message and continue with their other tasks. The interactive messages require the participant to record a message and also change information on patron's accounts. The reasoning messages involve library patrons calling to dispute a fine they received by mail. The participant is required to determine whether or not the patron should pay the fine by comparing the patron's demographic information in the master database to the information on the mailing list. The mailing list was used to send out fines; however the mailing list may have some errors. The participant is instructed that the master database has all of the correct information. If there is an error on the mailing list when compared to the master database, the participant should remove that patron's fine from their calculations on the inventory control sheet. There are two messages which fall under the reasoning category. In both reasoning messages, the patron was fined incorrectly due to an error in the mailing list. The participant should remove these fines from their calculations.

Prospective Memory

In addition to the tasks listed above, the participant is required to complete two prospective memory tasks: a time-based task and an event-based task. For the time-based task, the participant is asked to tell the examiner when ten minutes have passed from the start of the assessment. For the event-based task, the participant is instructed to notify the examiner when "Katie Hueil" calls. In addition, the participant is told to tell the examiner when they have finished the assessment. All instructions are provided at the beginning of the assessment and an instruction sheet is provided that the participant can view at any time during the assessment. Participants are instructed to ask any questions they have before the assessment, as they are not allowed to ask questions during the assessment. Participants have 40-minutes to complete the assessment; however they are not notified of the time limit before the assessment. Finally, participants are provided with a list of rules to follow during the assessment, e.g., do not talk to the examiner.

Scoring

The participant's total score is based on sub-scores from five categories based on the criteria established by Shallice and Burgess and used in the MET (Shallice & Burgess, 1991). Inefficiencies are observed when a more effective strategy could be used to complete the task. A rule break occurs when the participant has broken a specific rule that was given prior to the assessment or when a social rule was broken. An interpretation failure occurs when the participant misinterprets the requirements of the task. Task failures occur when a participant does not complete a task. If the task was not completed, the examiner determines the reason the participant did not complete the task. Accuracy on the inventory control task

is also recorded; however, it was not an original criterion in the MET. The total score of the CTPA is calculated by adding up the sub-score from each category. A score of 0 would indicate perfect performance while higher scores indicate a higher level of impairment.

Other Cognitive Assessments

The Weschler Test of Adult Reading (WTAR) is a measure of intellectual ability that asks the individual to read a list of irregularly spelled vocabulary words aloud; the WTAR was a measure also used to determine eligibility in this study. There is a well-established linear relationship between intelligence (IQ) and executive function performance (Bell & Roper, 1998; Diaz-Asper, Schretlen, & Pearlson, 2004; Jung, Yeo, Chiulli, Sibbitt, & Brooks, 2000; Timothy J. Wolf et al., 2010). If an individual has an impaired IQ score it is possible he/she will score poorly on a measure of executive function performance and in this case poor performance on a measure of executive function may actually be a reflection of lower overall intellectual capability. The WTAR is a valid and reliable measure of estimated IQ and has normative values for various populations (Strauss, Sherman, & Spreen, 2006). Inclusion criterion for this study was a score of greater than or equal to 85 which indicates average intelligence.

The Delis-Kaplan Executive Function System (DKEFS) is a nationally standardized measure of executive function battery. It consists of nine stand-alone subtests used to assess verbal and nonverbal executive function (Delis, Kramer, Kaplan, & Holdnack, 2004). For the purpose of this study, two subtests were used to examine the validity of the CTPA and determine eligibility for the study; Trail Making and Color-Word Interference. Trail Making requires the participant to sequence numbers and letters. Color-Word Interference requires a person to read a list of words based on the spelling or the color of the ink (Delis et al., 2004). These subtests demonstrate good reliability and validity (Delis, Kaplan, & Kramer, 2001; Delis et al., 2004).

Procedure

All testing took place at the Program in Occupational Therapy at Washington University in St. Louis. Informed consent was obtained upon arrival to the facility for testing. Once consent was obtained the participants completed the cognitive battery which includes the measures described above. Trained graduate occupational therapy students administered the assessment battery.

Community control subjects were tested at two different times. At initial testing CTPA performance was scored by the examiner and also video recorded and scored by an independent rater. The two test administrators alternated between being the primary live rater (directly administering and scoring the assessment) and the secondary rater (rating performance based on the video recordings). Participants also completed the cognitive assessment battery at the initial testing session. The person who was the primary live rater at the initial testing session also re-administered the CTPA at the second testing session one-week after initial testing. Mild stroke participants completed the same assessment battery but did not complete a second administration of the CTPA.

Analysis

Data were cleaned and entered into a RedCap © (Harris et al., 2009) database. All analysis was completed using SPSS 19.0. All data were plotted and evaluated for normality using the Shapiro-Wilk test. Several variables significantly deviated from a normal distribution so non-parametric statistical methods were used. Descriptive statistics were used to describe the populations and Mann-Whitney U tests (continuous data) or chi-squared analyses (categorical data) were used to evaluate differences between controls and individuals with stroke. Spearman rank correlation coefficients were calculated to evaluate the concurrent validity of the CTPA with the cognitive assessment battery and Mann-Whitney U tests with effect size calculations ($z/\sqrt{N}$) were used to evaluate the discriminant validity of the CTPA. Using data from control subjects only, intra-class correlation coefficients (ICC) were calculated to evaluate the inter-rater reliability and test-retest reliability on all scores sub-scores and total score of the CTPA.

Results

Participants

A description of participants is provided in Table 1 (N=34; 14 stroke, 20 controls). Participants in this study who sustained a mild stroke (n=14) were mostly female and African American. Community controls who participated in this study (n=20) were mostly female and Caucasian. Mann-Whitney U tests showed a significant difference between groups for age at testing, $z = -2.40$, $p = .016$ and WTAR standard score, $z = -3.80$, $p = .000$. Overall, community controls were younger ($M = 39.05$ SD =16.19) and obtained higher standard scores on the WTAR ($M = 112.15$ SD =9.29). Chi squared analysis revealed there were no differences between race, ($p = .110$), and gender, ($p = .868$), between groups. Finally, there was a significant difference between groups on the DKEFS subtests; however, this was expected based on the inclusion/exclusion criteria for this study.

Inter-rater and Test-retest Reliability

Intra-class correlation coefficients (ICCs) were calculated to determine inter-rater reliability of the CTPA total score and each of the sub-score categories. Results are shown in Table 2. Overall, the ICCs indicate a very high level of inter-rater reliability.

Intra-class correlation coefficients (ICCs) were calculated to determine test-retest reliability of the CTPA. The ICC for test-retest performance was 0.475 which did not meet the established $r=0.80$ level. A paired-samples t-test was also conducted to determine the difference in mean scores between each participant's first and second administration. There was a significant difference between administrations ($p = 0.006$) with a mean improvement of 5 points (11.4-6.4) from the first to second administration.

Concurrent Validity

Spearman rank correlation coefficients were calculated to determine the correlation between average CTPA total score performance and performance on the cognitive assessment battery. Results are shown in Table 3. Significant correlations were found between the CTPA and Conditions 4 the Color-Word Interference subtest which is considered the condition

specifically targeting executive function (Delis et al., 2001; Delis et al., 2004; Lippa & Davis, 2010). Significant correlations were also found with the Wechsler Test of Adult Reading. Overall, these results indicate an acceptable level of concurrent validity between the CTPA and gold standard neuropsychological measures of executive function.

Discriminant Validity

Mann-Whitney U tests found a significant difference between community controls and individuals with mild stroke on: (1) the total score of the CTPA, $z = -2.68$, $p = .007$; (2) total number of task failures, $z = -2.32$, $p = .020$; (3) total number of inefficiencies, $z = -2.01$, $p = .045$; and (4) total rule breaks, $z = -2.04$, $p = .042$. There was no significant difference on total items incorrect, $z = -1.78$, $p = .075$ and total interpretation failures, $z = -1.73$, $p = .083$. Overall, there were medium effects (r) on each score: (1) the total score of the CTPA, $r = -.460$; (2) total number of task failures, $r = -.340$; (3) total number of inefficiencies, $r = -.345$; (4) total rule breaks, $r = -.350$; (5) total items incorrect, $r = -.305$; and (6) total interpretation failures, $r = -.300$.

Discussion

This study examined the inter-rater, test-retest reliability, concurrent validity, and discriminant validity of the Complex Task Performance Assessment. Inter-rater reliability was excellent across all sub-scores and total score. The intra-class correlation coefficient did not meet the threshold ($r=0.80$) in order to have test-retest reliability. This may be due to a practice or learning effect of the assessment between test administrations which is expected with this form of assessment (Beglinger et al., 2004). These results support the notion that performance-based assessments should not be used as outcome measures in clinical practice or research intervention studies due to practice/learning effects. Test-retest reliability of most other performance-based assessments has not been thoroughly examined; however, it can be hypothesized that similar results would be expected with any performance-based assessment. Participants with mild neurological injury likely have intact episodic memory and their performance on the first administration is available for later introspection. Executive function is used to address novelty in our everyday life activities. As we gain experience with a task, laboratory testing has shown that frontal lobe activation/executive processing decreases (Doyon, Penhune, & Ungerleider, 2003; Meister et al., 2005; Puttemans, Wenderoth, & Swinnen, 2005). Additional work should be completed to establish alternate forms of these assessments that could be used as an outcome measure to evaluate changes in performance-based executive functioning.

The concurrent validity of the CTPA is acceptable as indicated by the correlations between the CTPA and cognitive assessments that target executive function. The correlation coefficients between the CTPA and the cognitive assessment battery are very similar to what have been found with other performance-based measures of executive function (Baum et al., 2008). These moderate correlations demonstrate a relationship between the CTPA and established measures of executive function but indicate that the assessments are uniquely different. This is evidenced by the moderate correlations with Condition 4 of the DKEFS Trail Making and Color-Word Interference subtests. Condition 4 of the Trail Making subtest

(Number-Letter Switching) is the condition meant to target executive function. The other conditions are baseline conditions meant to evaluate deficits in sequencing, visual scanning, and motor speed (Swanson, 2005; Yochim, Baldo, Nelson, & Delis, 2007). Condition 4 of the Color-World Interference subtest (Inhibition/Switching) was intended capture distinct aspects of executive functioning which could facilitate a diagnosis of executive function deficits that the other Color-Word Interference conditions cannot (Delis et al., 2001; Lippa & Davis, 2010). Further, the CTPA is significantly correlated with the WTAR total score (individuals who performed well on the WTAR also performed well on the CTPA), which also supports the concurrent validity of the CTPA. Individuals with higher intelligence (IQ) are more likely to have better performance on executive function assessments. The results from this study are consistent with this previously identified correlation between executive function performance and IQ (Bell & Roper, 1998; Diaz-Asper et al., 2004; Jung et al., 2000; Timothy J. Wolf et al., 2010).

Our results found that individuals with mild stroke performed significantly worse than community controls on the CTPA. The CTPA was able to discriminate between individuals with executive dysfunction and community controls on the CTPA total score and three of the subscores. The remaining two subscores were approaching significance and in a larger sample with normally distributed data set these differences may also be significant. These findings are consistent with previous research which has found that neuroperformance-based assessments, such as the MET, can discriminate between neurological impaired patients and controls (Alderman, Burgess, Knight, & Henman, 2003; Shallice & Burgess, 1991); however, the CTPA provides clinicians with an ecologically valid, performance-based measure which can detect these subtle deficits. Future work should focus on the evaluation of the sensitivity/specificity of the CTPA compared to other forms of assessment.

While our sample size was adequate to detect a difference between groups with statistical power, a larger sample size would further substantiate the discriminate validity of the CTPA. In addition, our sample was only representative of individuals with mild stroke. Executive dysfunction affects several diagnoses. Future testing should include other diagnoses known to experience deficits in executive function. Further, there was a relationship between CTPA total score and WTAR standard score. This correlation may suggest that there is an education effect where those with higher intelligence, as measured by the WTAR, perform better on the CTPA. While there was a relationship between CTPA total score and WTAR standard score, there was no relationship between CTPA total score and years of education between groups. Additional development should explore the relationship between education attainment, estimated intelligence, and executive function in more depth. In this study, the control group was younger, had a higher estimated IQ, and was predominately Caucasian. Any of these factors could have influenced the results and need to be accounted for statistically in future studies with the CTPA.

Currently, there are limited performance-based assessments that are able to detect mild deficits in executive function and are appropriate for use in the clinical setting. The CTPA provides clinicians with an ecologically valid, performance-based test that is sensitive to detecting executive dysfunction in the mild stroke population. A neuroperformance-based test, such as the CTPA, can be used to identify executive dysfunction and higher-level

cognitive deficits that are affecting performance. The CTPA only requires a few materials and a small amount of work space that could potentially allow for the generalizability of the assessment to be used in a variety of clinical settings. Performance based assessments such as the CTPA should not be used as outcome measures, but rather as a tool to identify performance-based executive function deficits. Future investigation should focus on the development of an alternate form so that the CTPA can be used as an outcome measure for treatment. The CTPA can provide therapists a way to identify performance-based executive function deficits in a clinical context before the person is allowed to return to family, community, and work life and fail.

Acknowledgements

The authors wish to acknowledge several individuals who have supported this study. First, Drs. Leonard Matheson and Traci Morrison who were involved in the initial development of the CTPA. Second, all of the students and staff in the Performance, Participation, and Neurorehabilitation Laboratory in the Program in Occupational Therapy at Washington University in St. Louis. Finally, we would like to specifically thank Kasey Porter, MSOT, Joanna Lackner, MSOT, and Jessica Seidman, MSOT.

Dr. Wolf was supported by the National Center for Medical Rehabilitation Research (NCMRR) in the National Institute of Child Health and Human Development (NICHD) of the National Institutes of Health under award numbers K12HD055931 and K23HD073190. This study was also supported by funding from the Boston Rehabilitation Outcomes Center funded by NCMRR as part of the Medical Rehabilitation Infrastructure Network (R24HD065688). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- Adams, RL., Parsons, OA., Culbertson, JL., Nixon, SJ. Neuropsychology for clinical practice: Etiology, assessment, and treatment of common neurological disorders. American Psychological Association; Washington D.C: 1996.
- Alaszewski A, Alaszewski H, Potter J, Penhale B. Working after a stroke: survivors' experiences and perceptions of barriers to and facilitators of the return to paid employment. *Disability & Rehabilitation*. 2007; 29; (24):1858–1869. [PubMed: 17852252]
- Alderman N, Burgess PW, Knight C, Henman C. Ecological validity of a simplified version of the multiple errands shopping test. *Journal of the International Neuropsychological Society*. 2003; 9; (1):31–44. [PubMed: 12570356]
- Baum CM, Connor LT, Morrison T, Hahn M, Dromerick AW, Edwards DF. Reliability, validity, and clinical utility of the Executive Function Performance Test: a measure of executive function in a sample of people with stroke. *American Journal of Occupational Therapy*. 2008; 62; (4):446–455. [PubMed: 18712007]
- Beglinger LJ, Gaydos B, Tangphao-Daniels O, Duff K, Kareken DA, Crawford J, Siemers ER. Practice effects and the use of alternate forms in serial neuropsychological testing. *Archives of Clinical Neuropsychology*. 2004; 20; (4):517–529.
- Bell BD, Roper BL. "Myths of neuropsychology": Another view. *Clinical Neuropsychologist*. 1998; 12; (2):237–244.
- Brott T, Adams HP, Olinger CP, Marler JR, Barsan WG, Biller J. Measurements of acute cerebral infarction: a clinical examination scale. *Stroke*. 1989; 20:864–870. [PubMed: 2749846]
- Delis, DC., Kaplan, E., Kramer, JH. Delis Kaplan Executive Function System. The Psychological Corporation; San Antonio, TX: 2001.
- Delis DC, Kramer JH, Kaplan E, Holdnack J. Reliability and validity of the Delis-Kaplan Executive Function System: an update. *Journal of the International Neuropsychological Society*. 2004; 10; (2): 301–303. [PubMed: 15012851]

- Diaz-Asper CM, Schretlen DJ, Pearlson GD. How well does IQ predict neuropsychological test performance in normal adults? *Journal of the International Neuropsychological Society*. 2004; 10; (1):82–90. [PubMed: 14751010]
- Doyon J, Penhune V, Ungerleider LG. Distinct contribution of the cortico-striatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia*. 2003; 41; (3):252–262. [PubMed: 12457751]
- Edwards D, Hahn M, Baum C, Dromerick A. The impact of mild stroke on meaningful activity and life satisfaction. *Journal of Stroke and Cerebrovascular Diseases*. 2006; 15; (4):151–157. [PubMed: 17904068]
- Edwards DF, Hahn MG, Baum CM, Perlmutter MS, Sheedy C, Dromerick AW. Screening patients with stroke for rehabilitation needs: Validation of the post-stroke rehabilitation guidelines. *Neurorehabilitation & Neural Repair*. 2006; 20; (1):42–48. [PubMed: 16467277]
- George M, Tong X, Kuklina E, Labarthe D. Trends in stroke hospitalizations and associated risk factors among children and young adults, 1995–2008. *Annals of Neurology*. 2011; 70; (5):713–721. [PubMed: 21898534]
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap) - A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics*. 2009; 42; (2):377–381. [PubMed: 18929686]
- Hartman-Maeir A, Soroker N, Ring H, Avni N, Katz N. Activities, participation and satisfaction one-year post stroke. *Disability & Rehabilitation*. 2007; 29; (7):559–566. [PubMed: 17453976]
- Hildebrand M, Brewer M, Wolf T. The impact of mild stroke on participation in physical fitness activities. *Stroke Research and Treatment*. 2012; 548682:6. Article ID.
- Jung RE, Yeo RA, Chiulli SJ, Sibbitt WL Jr, Brooks WM. Myths of neuropsychology: intelligence, neurometabolism, and cognitive ability. *Clinical Neuropsychologist*. 2000; 14; (4):535–545. [PubMed: 11262722]
- Lippa SM, Davis RN. Inhibition/switching is not necessarily harder than inhibition: an analysis of the D-KEFS color-word interference test. *Archives of Clinical Neuropsychology*. 2010; 25; (2):146–152. [PubMed: 20139109]
- Medin J, Barajas J, Ekberg K. Stroke patients' experiences of return to work. *Disability & Rehabilitation*. 2006; 28(17):10.
- Meister I, Krings T, Foltys H, Boroojerdi B, Muller M, Topper R, Thron A. Effects of long-term practice and task complexity in musicians and nonmusicians performing simple and complex motor tasks: implications for cortical motor organization. *Human Brain Mapping*. 2005; 25; (3): 345–352. [PubMed: 15852385]
- Norman, DA., Shallice, T. Attention to action: Willed and automatic control of behavior. In: Davidson, R.J., Schwartz, GE., Shapiro, D., editors. *Consciousness and Self-regulation*. Vol. 4. Plenum Press; 1986. p. 1-18.
- O'Brien A, Wolf TJ. Determining work outcomes in mild to moderate stroke survivors. *WORK*. 2010; 36; (4):441–447. [PubMed: 20714100]
- Puttemans V, Wenderoth N, Swinnen SP. Changes in brain activation during the acquisition of a multifrequency bimanual coordination task: from the cognitive stage to advanced levels of automaticity. *Journal of Neuroscience*. 2005; 25; (17):4270–4278. [PubMed: 15858053]
- Rochette A, Desrosiers J, Bravo G, St-Cyr-Tribble D, Bourget A. Changes in participation after a mild stroke: quantitative and qualitative perspectives. *Topics in Stroke Rehabilitation*. 2007; 14; (3):59–68. [PubMed: 17577968]
- Shallice T, Burgess PW. Deficits in strategy application following frontal lobe damage in man. *Brain*. 1991; 114:727–741. Pt 2. [PubMed: 2043945]
- Strauss, E., Sherman, E., Spreen, O. A compendium of neuropsychological tests: Administration, norms, and commentary. Oxford University Press; 2006.
- Swanson J. The Delis-Kaplan Executive Function System. *Canadian Journal of School Psychology*. 2005; 20(1-2):117–128.
- Wilson B, Alderman N, Burgess P, Ernsle H, Evans J. Behavioral Assessment of the Dysexecutive Syndrome. *Journal of Occupational Psychology, Employment and Disability*. 2003; 5; (2):33–37.

- Wolf T, Morrison M, Matheson L. Initial Development of a Work-Related Assessment of Dysexecutive Syndrome: the Complex Task Performance Assessment. *WORK*. 2008; 31; (2):221–228. [PubMed: 18957739]
- Wolf TJ, Barbee A, White D. Executive dysfunction immediately post-mild stroke. *OTJR*. 2010; 31(1) Supplement.
- Wolf TJ, Baum C, Connor LT. Changing face of stroke: Implications for occupational therapy practice. *The American Journal of Occupational Therapy*. 2009; 63; (5):621–625. [PubMed: 19785261]
- Wolfenden B, Grace M. Returning to work after stroke: a review. *International Journal of Rehabilitation Research*. 2009; 32; (2):93–97. [PubMed: 19158652]
- Yochim B, Baldo J, Nelson A, Delis D. D-KEFS Trail Making Test performance in patients with lateral prefrontal cortex lesions. *Journal of the International Neuropsychological Society*. 2007; 13; (04):704–709. [PubMed: 17521488]

Table 1

Description of Participants (N=34)

	Controls (n=20)	Stroke (n=14)
	Mean (Standard Deviation)	
Age *	39.05 (16.19)	52.93 (9.52)
Years of education	14.75 (1.86)	14.00 (2.35)
Wechsler Test of Adult Reading Standard Score **	112.15 (9.30)	97.14 (7.97)
DKEFS-Trail Making Condition 4-Scaled Score *	11.20 (2.48)	7.64 (3.52)
DKEFS-Color-Word Interference Condition 4-Scaled Score **	11.45 (2.31)	6.93 (3.36)
NIHSS Total	N/A	1.57 (1.45)
	Percentage	
Race (% Caucasian)	64%	29%
Sex (% female)	60%	57%

*
p<0.05;**
p<0.001

Table 2

Inter-rater Reliability of the Complex Task Performance Assessment (CTPA)

Score Category	ICC	95% CI	Mean (SD) Rater 1	Mean (SD) Rater 2
Total score	0.98	.94-.99	11.40 (8.76)	11.05 (8.82)
Inefficiencies	0.89	.74-.95	.50 (.69)	.50 (.69)
Rule Breaks	0.90	.77-.96	.55 (.83)	.50 (.95)
Interpretation Failures	0.95	.89-.98	2.55 (2.54)	2.30 (2.62)
Task Failures	0.95	.87-.98	1.10 (1.21)	1.15 (1.18)

Table 3

Correlations of the Complex Task Performance Assessment (CTPA) to the Cognitive Assessment Battery

Neuropsychological Assessment	Correlation to CTPA total score average (ρ)	p
DKEFS Trail Making Condition 4	-0.24	0.171
DKEFS Color-Word Condition 4	-0.43 *	0.012
Wechsler Test of Adult Reading	-0.49 **	0.003

Note: DKEFS indicates Delis-Kaplan Executive Function Systems Assessment; Color-Word indicates the Color-Word Interference subtest of the DKEFS. ρ = Spearman rank correlation coefficient .

* indicates significant at the 0.05 level (2 tailed);

** indicates significant correlation at the 0.01 level (2-tailed).