Cognitive Correlates of Functional Performance in Older Adults: Comparison of Self-Report, Direct Observation, and Performance-Based Measures

Maureen Schmitter-Edgecombe,[1] Carolyn Parsey,[1] AND Diane J. Cook[2]
[1]Department of Psychology, Washington State University, Pullman, Washington
[2]School of Electrical Engineering and Computer Science, Washington State University, Pullman, Washington

Abstract

Neuropsychologists are often asked to answer questions about the effects of cognitive deficits on everyday functioning. This study examined the relationship between and the cognitive correlates of self-report, performance-based, and direct observation measures commonly used as proxy measures for everyday functioning. Participants were 88 community-dwelling, cognitively healthy older adults (age 50–86 years). Participants completed standardized neuropsychological tests and questionnaires, and performed eight activities of daily living (e.g., water plants, fill a medication dispenser) while under direct observation in a campus apartment. All proxy measures of everyday function were sensitive to the effects of healthy cognitive aging. After controlling for age, cognitive predictors explained a unique amount of the variance for only the performance-based behavioral simulation measure (i.e., Revised Observed Tasks of Daily Living). The self-report instrumental activities of daily living (IADL) and the performance-based everyday problem-solving test (i.e., EPT) did not correlate with each other; however, both were unique predictors of the direct observation measure. These findings suggest that neuropsychologists must be cautious in making predictions about the quality of everyday activity completion in cognitively healthy older adults from specific cognitive functions. The findings further suggest that a self-report of IADLs and the performance-based EPT may be useful measures for assessing everyday functional status in cognitively healthy older adults. (*JINS*, 2011, *17*, 853–864)

Keywords: Activities of daily living, Cognitive correlates, Functional status, Aging, Instrumental activities of daily living, Everyday functioning

INTRODUCTION

To live independently at home, individuals need to be able to complete everyday instrumental activities of daily living (IADL), including preparing meals, managing money and taking medications. Due to the aging of the population, the number of individuals unable to live independently in their homes because of cognitive impairments is rising rapidly. Consequently, neuropsychologists are increasingly being asked to answer questions regarding the effects of cognitive deficits on everyday functioning (Marcotte, Scott, Kamat, & Heaton, 2010). Currently, however, there is no "gold standard" for measuring everyday functional abilities. In addition, the approach of using neuropsychological tests originally designed for other purposes (e.g., documentation of brain injury severity, localization of brain pathology) to predict everyday functioning has been questioned (Goldstein, 1996).

Self-report, informant-report, and performance-based measures have commonly been used in the literature as a proxy for real-world functioning. Each method has distinct advantages and disadvantages. For example, while self-report and informant-report questionnaires are easy to administer and may give a reasonably accurate representation of real-world performance given the opportunity for multiple observations, they are subject to reporter bias (Bertrand & Willis, 1999; Dassel & Schmitt, 2008; Richardson, Nadler, & Malloy, 1995). Similarly, while performance-based measures are objective, quantifiable, repeatable, and norm-referenced, the role of the environment is typically subtracted and performance on such measures can fluctuate with motivation, cognition and behavior (Marson & Hebert, 2006; Myers, Holliday, Harvey, & Hutchinson, 1993; Zimmerman & Magaziner, 1994).

Correspondence and reprint requests to: Maureen Schmitter-Edgecombe, Department of Psychology, Washington State University, P.O. Box 644820, Pullman, WA 99164-4820. E-mail: schmitter-e@wsu.edu.

Two types of performance-based tests have been developed. Paper and pencil tasks assess everyday cognition or everyday problem-solving by presenting cognitively challenging real-world problems (e.g., Everyday Problems Test [EPT]; Willis & Marsiske, 1993; Everyday Cognition Battery; Allaire & Marsiske, 1999, 2002), while behavioral simulations tasks require individuals to complete everyday tasks within a laboratory environment (e.g., write a check, look up a phone number; Revised Observed Tasks of Daily Living [OTDL-R], Diehl et al., 2005; Texas Functional Living Scale, Cullum et al., 2001). Currently, the question of which method for assessing functional performance most closely approximates actual behavior in the real world remains debatable (Farias, Harrell, Neumann, & Houtz, 2003).

It has been argued that direct observation of individuals in the everyday environment will provide the most valid determination of everyday functional status (Marcotte et al., 2010), as it would allow for observation of behaviors with increased subtlety and across extended periods of time. New technologies, which use sensors and machine learning algorithms to identify and track activities of daily living continuously within the home environment, are being developed (Hayes et al., 2008; Rashidi, Cook, Holder, & Schmitter-Edgecombe, 2011). In this study, we report on behavioral observation data collected while healthy older adults completed activities of daily living within a campus apartment. The apartment had a kitchen, dining room, and living room on the downstairs level. The behaviors that were observed included eight, highly scripted complex everyday activities (e.g., locate and fill a medication dispenser; prepare a cup of noodle soup) that were each performed once by the participant. While the on-campus apartment presented a unique opportunity to collect direct behavioral observation data within an everyday environment, participants were not tested in their own homes nor were repeated measurements of the same everyday activities collected.

One goal of the present study was to examine the relationships between several methods for assessing functional status in a healthy older adult population, including direct behavioral observation, self-report, and performance-based measures. Prior research suggests that self-report measures are generally accurate indicators of daily functioning for older adults who demonstrate insight into their functional abilities (Alexander et al., 2000; Farias, Mungas & Jagust, 2005). The performance-based tasks included a measure of everyday problem-solving (i.e., EPT) and a behavioral simulation measure (i.e., OTDL-R). Only a few studies have directly compared more than one method for assessing functional status. These studies have generally found that different methods for assessing functional status do not correlate highly with each other and further can provide different estimates of an individual's ability to perform everyday activities (e.g., Burton et al., 2009; Jefferson et al., 2008, Loewenstein et al., 2001). Given the significant implications that poor functional status can have for independent living, it is important to better understand the relationship amongst these varied methods used to assess functional status and, ultimately, to understand the ability of such measures to serve as a surrogate for real-world, everyday functioning.

A second goal of the study was to examine which cognitive correlates, as measured by neuropsychological tests, are most predictive of quality of everyday activity completion across the different methods used. Understanding the cognitive determinants of performance on these proxy measures of everyday functioning may improve our knowledge of what aspects of everyday performance these diverse measures of functional status are capturing. Prior research suggests that there are many cognitive factors that may contribute to functional impairment in older adult populations, including global cognitive functioning (Farias et al., 2003; Inzarti & Basile, 2003), memory (Farias et al., 2006; McCue, Rogers, & Goldstein, 1990), processing speed (Tuokko, Morris, & Ebert, 2005), visuo-perceptual abilities (Glosser et al., 2002; Jefferson, Barakat, Giovannetti, Paul, & Glosser, 2006), and executive functioning (Bell-McGinty, Podell, Franzen, Baird, & Williams, 2002; Cahn-Weiner, Malloy, Boyle, Marran, & Salloway, 2000; Lewis & Miller, 2007). Measures of executive functioning have, however, emerged across studies as the strongest and most consistent specific cognitive predictor of everyday functional status (Cahn-Weiner et al., 2007; Marcotte et al., 2010).

Numerous studies have also provided evidence for non-cognitive risk factors of functional impairment in older adult populations, including advanced age, lower education, depressed mood, and comorbid health conditions (e.g., knee replacement). For example, depression has been found to be a significant predictor of functional disability in some studies (Cho et al., 1998; Guccione et al., 1994) but not others (Galanos, Fillenbaum, Cohen, & Burchett, 1994). Therefore, when examining the relationship between cognition and functional status, it is important to understand how much variance in functional status can be attributed to cognition independent of major non-cognitive covariates.

In this study, healthy community-dwelling older adults completed self-report and performance-based measures of everyday functioning in a laboratory environment. Observational assessment of participant's ability to complete everyday activities was conducted in a university campus apartment. Given prior research findings (e.g., Jefferson et al., 2008; Rueda & Schmitter-Edgecombe, 2011), we did not expect strong correlations between the self-report and performance-based measures. We were especially interested in how performance on the direct observation measure would relate to both the self-report and performance-based measures, which are often used as proxy measures for everyday functional status. Based on prior research (e.g., Cahn-Weiner et al., 2000), we also hypothesized that, after controlling for demographic variables, the neuropsychological measure of executive functioning would emerge as the strongest predictor of quality of everyday activity completion in cognitively healthy older adults.

METHOD

Participants and Procedure

Participants were 88 community-dwelling healthy older adults. To assess whether limitations in functional status are

more common in individuals age 75+ (Lafortune & Balestat, 2007) rather than also being apparent in the young-old, we initially differentiated the healthy older adults into three age groups, middle-aged ($N = 22$; age 50–59 years; 18 F, 4 M), young-old ($N = 44$; age 60–74 years; 29 F, 15 M) and old-old ($N = 22$; age 75–86 years; 16 F, 6 M). Exclusionary criteria included history of head trauma with permanent brain lesion, history of cerebrovascular accidents, current or recent (past year) psychoactive substance abuse, known medical, neurological, or psychiatric causes of cognitive dysfunction (e.g., schizophrenia, epilepsy), and self- or knowledgeable informant report of significant memory complaints or changes in cognitive ability across the past months to years. Initial screening for cognitively healthy older adult participants was conducted over the phone and included: (a) a medical interview to rule out exclusion criteria, (b) the Telephone Interview for Cognitive Status (TICS) to exclude participants who scored below 27 (equivalent of an MMSE of 24) on a measure of global cognitive functioning (Brandt & Folstein, 2003), and (c) the Clinical Dementia Rating instrument (CDR) to rule out cognitive impairment suggestive of questionable dementia (i.e., CDR > 0; Hughes, Berg, Danzinger, Coben, & Martin, 1982; Morris, 1993). Healthy older adult participants were recruited through community advertisements, community health and wellness fairs, and presentations to community groups consisting of a large numbers of seniors.

Participants who met initial screening criteria completed a laboratory battery of standardized and experimental neuropsychological tests (approximately 3 hr). To rule out participants who might be experiencing mild cognitive impairment, participants were psychometrically determined to be cognitively healthy if they performed within 1.5 standard deviations of age-corrected and education-corrected (when available) standardized test norms on the cognitive tests in Table 1. These participants then completed a variety of complex activities of daily living within an apartment located on the Washington State University campus (e.g., sweep the floor, water plants). This evaluation took approximately 3 hours. Participants were given a report documenting their performance on the neuropsychological tests, as well as pre-paid parking passes, as compensation for their time. Participants who traveled to the laboratory from outside Whitman or Latah County were also provided a $50 voucher for travel reimbursement. This protocol was reviewed and approved by the Institutional Review Board at WSU.

Table 1. Demographic data and mean summary data for middle-aged, young-old, and old-old groups

	Group					
	Middle-aged ($n = 22$)		Young-old ($n = 44$)		Old-old ($n = 22$)	
Variable or test	M	SD	M	SD	M	SD
Demographics						
Age	54.55	3.36	67.14	4.41	80.00	3.06
Education (years)	16.95	2.15	16.81	2.92	16.77	2.49
GDS	1.77	2.16	1.63	1.91	1.62	1.24
Global cognitive status						
Shipley total score	34.23	4.05	35.36	2.98	36.36	2.19
TICS total score	34.41	2.38	34.98	2.49	33.37	2.80
Attention/Processing Speed						
SDMT Oral total	61.73	8.09	59.32	9.96	47.41[ab]	10.32
SDMT Written total	53.41	6.92	51.16	9.60	42.05[ab]	8.00
Verbal Memory						
MAS list learning	62.68	5.63	60.61	5.40	59.36	6.47
MAS Delayed Recall[†]	11.41	1.01	11.53	.73	11.10	1.38
Word Finding						
BNT total correct	56.77	2.25	57.43	2.34	56.68	2.85
Executive Skills						
WAIS-III L-N Seq.	10.77	2.54	11.00	2.47	9.22[b]	2.69
D-KEFS Letter Fluency	41.23	12.63	43.91	9.32	44.73	11.35
D-KEFS Design Fluency	28.82	6.12	28.48	6.74	24.91	7.48
Visuoperception						
CLOX copy	13.64	1.05	13.45	1.02	13.41	1.04

Note. Unless otherwise indicated, mean scores are raw scores. Norm sources for the cognitive tests are in parentheses following the test. GDS = Geriatric Depression Scale; TICS = Telephone Interview for Cognitive Status (Brandt & Folstein, 2003); SDMT = Symbol Digit Modalities Test (Smith, 1991); MAS = Memory Assessment Scale (Williams, 1991); BNT = Boston Naming Test (Ivnik, Malec, Smith, Tangalos, & Petersen, 1996); WAIS-III L-N Seq. = Letter–Number Sequencing subtest of the Wechsler Adult Intelligence Scale—Third Edition (Wechsler, 1997); D-KEFS = Delis-Kaplan Executive Function System (Delis, Kaplan, & Kramer, 2001).
[a]Significant difference compared with middle-aged group, $p < .05$.
[b]Significant difference compared with young-old group, $p < .05$.

Measures

Non-cognitive risk factors

Non-cognitive risk factors included participants' age, level of education, and number of depressive symptoms. The total number of symptoms of depression reported on the Geriatric Depression Scale – Short Form (Yesavage et al., 1983) was used as the measure of depressive symptomology.

Cognitive test variables

The cognitive predictor variables were derived from standardized neuropsychological tests administered during the laboratory assessment and represent the following cognitive constructs: global cognitive status, processing speed, memory, visuoperceptual abilities, and executive functioning. A brief description of each task can be found in Appendix A; below we indicate the score from each task that was used as the cognitive predictor variable.

Telephone Interview for Cognitive Status (TICS; Brandt & Folstein, 2003)

The total score from this mental status exam was used as the measure of global cognitive functioning.

Symbol Digits Modalities Test (SDMT; Smith, 1991)

To reduce the impact that health-related variables, such as arthritis, might have on the data, the total score from the oral version of the SDMT was used as the measure of processing speed.

Memory Assessment Scale (MAS) List Learning subtest (Williams, 1991)

The number of 12 words recalled at the long-delay was used as the measure of memory.

Clox 1&2 (Royall, Cordes, & Polk, 1998)

Total score on the copy subtask of the clock (Clox 2) was used as the measure of visuoperceptual abilities.

Delis-Kaplan Executive Function System (D-KEFS) Letter Fluency subtest (Delis, Kaplan & Kramer, 2001)

The total number of correct words produced for the letters F, A, and S was used as the measure of executive functioning.

Functional status measures

The domains of IADLs that were assessed across the self-report, performance-based, and direct observation measures were largely consistent (e.g., medication management, meal preparation). For the purposes of this study, these four different types of measures that have been used in the literature as a proxy for real-world functioning, are labeled collectively as functional status measures.

Self-report IADL

Participants completed the self-report Lawton-Brody IADL scale (Lawton, Moss, Fulcomer, & Kleban, 1982), indexing nine IADL domains: using the phone, traveling, shopping, preparing meals, doing housework, handyman work, laundry, medication management, and financial management. Participants rated their skill level (capacity) for each IADL domain using a Likert scale, ranging from 1 (completely unable to do) to 3 (can complete without help). The items were summed to provide a total score.

The Revised Observed Test of Daily Living

The OTDL-R (Diehl et al., 2005) is a performance-based measure of everyday competency. Participants are presented with problem scenarios and real-life materials (e.g., medicine bottles) and asked to perform the necessary steps to find the correct answer to questions presented on cards (e.g., look at three medicine bottle labels and indicate how many days a specific refill will last). The OTDL-R includes a total of nine IADL tasks (28 items, maximum score = 28), three each in the domains of medication use, telephone use, and financial management.

The Everyday Problems Test

The EPT (Willis & Marsiske, 1993) is a paper and pencil performance-based measure of everyday cognitive competence. Participants are required to solve problems of daily living using printed real-life tasks drawn from seven domains of functioning stimuli (e.g., examine a chart of taxi rates and choose how much they would have to pay to travel one mile in a suburban area from four responses). For the purposes of this study, four domains were assessed (i.e., shopping, transportation, household, meal preparation), as the remaining three domains were covered by the OTDL-R. The items were summed to obtain a total score (maximum score = 24), with each correct item receiving a score of 1.

Direct observation of everyday activity completion

Participants completed eight activities of daily living within a campus apartment (e.g., filling a 7-day medication dispenser with three types of medications). See Appendix A for a more detailed description of each of the activities. Before beginning each activity, experimenters provided brief verbal instructions and asked that all materials be returned to their original positions after task completion. Participants then carried out each activity using the materials provided to them within the apartment. As the participant completed the activities, two experimenters observed the participant and coded the actions based on the sequence and accuracy of the steps completed. The experimenters also recorded extraneous participant actions (e.g., searching for items in wrong locations). Each activity was coded for six different types of errors: critical omissions, critical substitutions, non-critical omissions, non-critical substitutions, irrelevant actions and

Table 2. Coding schema used to derive the direct observation score and scoring reliability

Each activity was coded for instances of the following six error types. An overall direct observation score was then derived for each activity.

Critical Error Types

1. *Critical Omission*: Coded when a step or subtask that is necessary for accurate completion of the activity is not completed (e.g., failure to retrieve broom, failure to put check into envelope).
2. *Critical Substitution*: Coded when an alternate object, or a correct object but an incorrect gesture, is used and disrupts accurate completion of the activity (e.g., dusting the kitchen instead of the living room, filling medication dispenser incorrectly).

Non-Critical Error Types

1. *Non-Critical Omission*: Coded when a step or subtask is not performed but the activity is still completed accurately (e.g., failure to return items to their original locations, does not turn off the television).
2. *Non-Critical Substitution*: Coded when an alternate object, or a correct object but an incorrect gesture, is used but the activity is still completed accurately (e.g., uses buttons on television rather than remote to operate electronics, uses container other than watering can to water plants).

Extraneous Error Types

1. *Irrelevant Action*: Coded when an action that is unrelated to the activity, and completely unnecessary for activity completion, is performed (e.g., sweeping the front entryway in addition to the kitchen, washes dishes after completing the cooking task).
2. *Inefficient Action*: Coded when an action that slows down, or compromises the efficiency of task completion, is performed (e.g., making multiple trips to the dining room table, opening and closing extraneous cupboards and drawers unnecessary for task completion).

Direct Observation Score

  1 = task completed without any errors
  2 = task completed but with no more than 2 total of the following errors: non-critical omissions, non-critical substitutions, irrelevant actions, inefficient actions
  3 = task completed but with more than 2 total of the following errors: non-critical omissions, non-critical substitutions, irrelevant actions, inefficient actions
  4 = task incomplete, coded when a <u>critical</u> omission or substitution occurs; more than 50% of the task must be completed
  5 = task incomplete, coded when a <u>critical</u> omission or substitution occurs; less than 50% of the task completed

Scoring reliability

The recorded data was coded by two scorers working independently. The raters were blind to the age of the individual. Agreement was 96.49% for critical errors, 96.84% for noncritical errors, and 97.79% for extraneous errors. In addition, agreement was 95.45% for the overall activity score. A list of potential errors was generated for each error category for each activity, and when new errors were encountered they were added to the list. When discrepancies arose, the two scorers discussed the error in question and resolved the discrepancy. The coders had access to videotapes of each participant completing the activities.

inefficient actions. Table 2 provides detailed code assignment information for each error as well as the scoring rubric used to derive an overall score for each activity and scoring reliability. The overall score for each of the eight activities was summed to derive the direct observation score.

RESULTS

Demographic and Neuropsychological Data

Table 1 shows the demographic and neuropsychological testing data by age group. One-way analyses of variance (ANOVAs) followed by post hoc contrasts using Tukey's HSD (honestly significant difference) revealed that the old-old performed more poorly than the young-old on a test measuring working memory/executive abilities, $F(2,85) = 3.73$; $p < .05$, and more poorly than both the middle-aged and young-old on the written, $F(2,85) = 11.30$, $p < .001$, and oral, $F(2,85) = 14.84$, $p < .001$, trials of a test measuring attention and speeded processing (see Table 1). The age groups did not differ significantly in level of education, self-report of depressive symptoms, or their performances on the remainder of neuropsychological measures displayed in Table 1.

Functional Status Measures

One-way ANOVAs followed by post hoc contrasts were then used to examine whether the mean scores for the middle-aged, young-old, and old-old differed across the four measures of functional status.[1] Because there were no gender differences in performance on the functional status measures, t's > 1.10, p's < .30, sex was not included as a variable in the analyses. Table 3 displays the means and standard deviations for each functional status measure by age group. Significant group differences were found for all functional status measures, F's > 3.70, p's < .05. As seen in Table 3, the old-old group showed poorer performance than the middle-aged group on all functional status measures, including self-report IADL ($d = .92$),[2] OTDL-R ($d = 1.31$), EPT ($d = .74$), and direct observation ($d = 1.72$). The old-old group also differed from the young-old group on self-report IADL ($d = .86$),

---

[1] Because a ceiling effect impacted the normality of the distribution for the self-report IADL measure for all three age groups, we also assessed for group differences in the self-report IADL measure with the Kruskal-Wallis non-parametric test followed by *post hoc* Mann Whitney *U* tests. The findings were identical to those reported using the parametric statistics.

[2] Effect sizes reported in parentheses represent Cohen's *d*.

Table 3. Mean summary data for middle-aged, young-old, and old-old groups on the functional status measures

	Group					
	Middle-aged ($n = 22$)		Young-old ($n = 44$)		Old-old ($n = 22$)	
Functional status Measure	M	SD	M	SD	M	SD
Self-report IADL	26.36	.66	26.45	.95	25.67[ab†]	.86
EPT	22.32	1.32	21.69[§]	1.79	21.00[a‡]	2.28
OTDL-R	24.43[†]	2.23	23.05[‡‡]	3.09	20.59[ab]	3.46
Direct observation	11.18	2.15	13.55[a]	3.31	16.45[ab]	3.77

Note. Mean scores are raw scores. IADL = Instrumental Activities of Daily Living; EPT = Everyday Problem-solving; OTDL-R = Revised Observed Test of Daily Living. [†]n = 21; [‡]n = 18; [§]n = 39; [‡‡]n = 43.
[a]Significant difference compared with middle-aged group, $p < .05$.
[b]Significant difference compared with young-old group, $p < .05$.

OTDL-R ($d = .76$) and direct observation ($d = .84$) measures. In addition, the direct observation score of the young-old group was poorer than that of the middle-aged group ($d = .80$). These data indicate that all four proxy measures of functional status revealed greater functional limitations in the old-old group.

Intercorrelations between the Measures of Functional Status

Table 4 shows the intercorrelations amongst the four measures of functional status with age partialled out. No significant correlations were found between the self-report IADL and two performance-based measures (i.e., OTDL-R and EPT). The correlation between the OTDL-R and EPT just failed to reach statistical significance ($r = .21$; $p = .066$). The direct observation measure correlated with both the self-report IADL ($r = -.31$) and EPT ($r = -.49$), indicating that those with poorer direct observation scores tended to self-report more everyday difficulties and performed more poorly on the EPT.

Regression Analyses: Cognitive Determinants

Hierarchical regression analyses were completed to identify cognitive determinants of everyday functional status after controlling for non-cognitive risk factors. To reduce the number of predictor variables included in the regression analyses, we first examined the correlations between the four measures of functional status and the non-cognitive risk

factors. All functional outcome measures correlated significantly with age but not with education or symptoms of depression (see Table 5). Subsequently, only age was entered in as non-cognitive risk factors in the first block of the hierarchical regression. In block 2, we entered the cognitive predictors derived from the neuropsychological testing data. The cognitive predictors were chosen to represent different domains of cognitive abilities that have been found to be predictive of functional status in prior studies: TICS total score (global cognitive functioning), MAS delayed total recalled (memory), SDMT oral total correct (processing speed), Clox 2: copy total score (visuoperceptual abilities), and D-KEFS letter fluency (executive functioning). The cognitive predictors were entered simultaneously in the second step to determine if they held any unique and predictive value. There was no significant multicollinearity amongst the five cognitive predictor variables, as the Variance Inflation Factors for each variable were less than 1.6. Correlations amongst the predictor and criterion variables can be found in Table 5.

The cognitive predictors (variance accounted for represented by ΔR^2) were found to account for significant variance over and above that accounted for by age for only the performance-based OTDL-R task [$\Delta R^2 = .17$; $\Delta F(5,75) = 4.41$; $p = .001$; total $R^2 = .42$]. For the OTDL-R, the SDMT-oral was a unique cognitive predictor, $B = .31$, $t = 2.83$, $p = .006$. The neuropsychological measures did not account for significant variance over and above age for either the self-report IADL [$\Delta R^2 = .08$; $\Delta F(5,76) = 1.36$; $p = .25$;

Table 4. Correlation matrix for the four functional status measures (after controlling for age)

	Self-reported IADL	EPT	OTDL-R	Direct observation
Self-report IADL	1.0	.03	−.05	−.31*
EPT		1.0	.21	−.49**
OTDL-R			1.0	−.16
Direct observation				1.0

Note. IADL = Instrumental Activities of Daily Living; EPT = Everyday Problems Test; OTDL-R = Revised Observed Test of Daily Living.
*$p < .005$; ** $p < .001$.

Table 5. Correlations between the functional status measures and Non-Cognitive Risk Factors and Neuropsychological Testing Data.

	Functional status measures			
Variables	Self-report IADL	OTDL-R	EPT	Direct observation
Age	$-.21*$	$-.49**$	$-.34**$	$.53**$
Education	.16	.07	.02	$-.16$
GDS: depression	$-.18$	$-.03$	.11	$-.01$
TICS: Global Cognitive Status	.21	$.28*$	$.25*$	$-.26*$
SDMT oral: Processing Speed	$.25*$	$.49**$	.11	$-.35**$
MAS delayed recall: Memory	$-.08$	.19	.10	$-.17$
D-KEFS LF: Executive Functioning	$-.02$	.20	$-.08$	$-.07$
Clox 2, copy: Visuoperception	.04	.17	.21	$-.24*$

Note. Total correct raw score was used for all neuropsychological measures. IADL = Instrumental Activities of Daily Living; EPT = Everyday Problems Test; OTDL-R = Revised Observed Test of Daily Living; GDS = Geriatric Depression Scale; TICS = Telephone Interview for Cognitive Status; SDMT = Symbol Digit Modalities Test; MAS = Memory Assessment Scale; D-KEFS Letter Fluency subtest.
$*p < .05$; $**p < .005$.

total $R^2 = .13$], EPT [$\triangle R^2 = .06$; $\triangle F(5,68) = 1.06$; $p = .39$; total $R^2 = .18$] or direct observation score [$\triangle R^2 = .07$; $\triangle F(5,77) = 1.60$; $p = .17$; total $R^2 = .35$]. In addition, there were no significant unique cognitive predictors for the self-report IADL, t's < 1.78, p's $> .08$, EPT, t's < 1.55, p's $> .12$, and direct observation, t's < 1.60, p's $> .10$, measures. These findings suggest that while the cognitive variables explained a unique amount of the variance in the performance-based OTDL-R measure, the cognitive predictors did not explain unique variance above that of age for the self-report IADL, EPT, and direct observation measures.[3]

Regression Analyses: Functional Status Proxy Measures

Given that the cognitive determinants did not predict performance on the direct observation measure, we were interested in whether the self-report and performance-based measures, which are considered proxy measures for everyday functioning, would contribute unique variance in predicting the direct observation score. Age was again entered in the first step of the hierarchical regression, followed by the functional status measures (self-report IADL, OTDL-R, EPT) in the second step. The results of this regression analysis revealed that the proxy measures of everyday functioning accounted for significant variance over and above that accounted for by age [$\triangle R^2 = .21$; $\triangle F(3,72) = 9.84$; $p < .001$], with both the measures of everyday functioning and age accounting for 47% of

---

[3] The regression analyses were also run excluding age as a predictor. The findings similarly revealed that the cognitive predictors did not account for significant variance in the self-report IADL [$R^2 = .11$; $F(5,77) = 1.99$; $p = .09$] or the EPT [$R^2 = .11$; $F(5,69) = 1.68$; $p = .15$]. The cognitive predictors continued to account for a significant amount of the variance in the OTDL-R [$R^2 = .35$; $F(5,76) = 8.19$; $p < .001$], with both the SDMT-oral, $B = .47$; $t = 4.92$; $p < .001$, and the Clox 2 (copy) serving as unique cognitive predictors, $B = .19$; $t = 2.00$; $p < .05$. The cognitive predictors now also accounted for a significant amount of the variance in the direct observation score [$R^2 = .35$; $F(5,76) = 8.19$; $p < .001$], with the SDMT-oral, $B = -.36$, $t = -3.44$, $p = .001$, and the Clox 2 (copy) both serving as unique cognitive predictors, $B = -.22$, $t = -2.16$, $p < .05$.

the variance in the direct observation measure, $R^2 = .47$. In addition, both the EPT, $B = -.37$, $t = -4.07$, $p < .001$, and self-report IADL, $B = -.25$, $t = -2.94$, $p = .004$, were unique predictors. If we make the assumption that the direct observation measure is the better proxy for everyday functional performance, the data suggest that the everyday problem-solving measure and the self-report IADL measure are better at predicting everyday functioning in this group of cognitively healthy older adults than the cognitive predictors and the laboratory behavioral simulation measure (i.e., the OTDL-R).

In further support, when we looked at the ability of the self-report IADL, EPT and direct observation scores to predict OTDL-R performance over and above age, we found that no significant additional variance was accounted for by the functional status measures [$\triangle R^2 = .03$; $\triangle F(3,72) = 1.15$; $p = 34$; total $R^2 = .30$]. This finding contrasts sharply with the 17% additional variance accounted for by the cognitive variables when predicting OTDL-R performance. This suggests that the OTDL-R may be associated more closely with similar abilities involved in other cognitive tests rather than with functional status measures. Regression analyses for both the EPT [$\triangle R^2 = .19$; $\triangle F(3,72) = 6.51$; $p = .001$; total $R^2 = .30$] and the self-report IADL [$\triangle R^2 = .12$; $\triangle F(3,72) = 3.07$; $p = .03$; total $R^2 = .16$] showed that the functional status measures accounted for significant variance over and above that of age. In both analyses, the direct observation measure was found to be a unique predictor (EPT, $B = -.57$; $t = -4.62$; $p < .001$; self-report IADL, $B = -.51$; $t = -4.07$; $p < .001$). In addition, the amount of variance accounted for by age and the functional status measures was greater than that accounted for by age and the cognitive predictors for both the self-report IADL (16% *vs.* 13%) and EPT (30% *vs.* 18%).[4]

---

[4] The regression analyses were also run excluding age as a predictor. The analysis similarly revealed that the amount of variance accounted for by the functional status measures was significant for the self-report IADL [$R^2 = .15$; $F(3,73) = 4.28$; $p = .008$] and the EPT [$R^2 = .30$; $F(3,73) = 10.24$; $p < .001$]. The direct observation measure also continued to serve as a unique predictor for both the self-report IADL, $B = -.52$, $t = -4.49$, $p < .001$, and

DISCUSSION

Neuropsychologists routinely use data collected from neuropsychological tests to predict real-world functioning. Evaluation of the ability of traditional neuropsychological tests to predict functional status has, however, been hampered by lack of a "gold standard" measure for the assessment of everyday functional abilities. We compared four types of measures that have been used in the literature as a proxy for real-world functioning, including self-report, performance-based measures and direct observational data in a real-world setting. We also sought to identify the cognitive correlates of these functional status measures after controlling for non-cognitive risk factors.

All four proxy measures showed sensitivity to the healthy aging process. In addition, all four measures differentiated the old-old and middle-age groups, and all but the EPT differentiated the old-old and young-old groups. Only the direct observation measure differentiated the performances of the middle-aged and young-old group. These findings are consistent with prior research (Lafortune & Balestat, 2007), in suggesting that individuals age 75+ are at greater risk for limitations in everyday functioning. Of note, 47% of the older adults endorsed full independence on all self-report IADL domains assessed, limiting the sensitivity of this measure when used with a cognitively healthy aging population.

When the intercorrelations amongst the functional status measures were compared (controlling for age), the correlation between the two performance-based measures just failed to reach significance. This may partially reflect the fact that the two performance-based measures assessed slightly different subsets of IADL domains and represent different ways of assessing functional status. Although both measures make use of stimuli that individuals are exposed to in everyday life, the OTDL-R is a behavioral simulation measure while the EPT primarily taps into everyday problem-solving abilities. The correlational data further revealed that while the self-report IADL and EPT measures did not correlate with each other, both measures correlated with the direct observation measure. Lack of correlation between self-report and performance-based measures has been found in prior studies when these two measures have been compared (e.g., Kempen, Steverink, Ormel, & Deeg, 1996; Reuben, Valle, Hays, & Siu, 1995). This has led to the question of whether self/informant reports or performance-based measures are a better representation of an individual's actual real-world behaviors. If data from the direct observation measure is considered to be the superior measure of everyday functional status, this may

suggest that the self-report IADL and EPT may be measuring different aspects of everyday performance in this cognitively healthy aging population. For example, while the self-report IADL might be tapping into knowledge gained from multiple experiences completing everyday activities and other physical health variables, the EPT might be tapping into the ability to use and apply everyday problem-solving skills. The possibility that these two measures are evaluating different aspects of everyday functioning was further supported by the results of the regression analysis, which showed that both the self-report IADL and EPT contributed unique variance to the prediction of participants' direct observation scores.

For the self-report IADL, EPT and direct observation measures, after controlling for age, the cognitive variables did not explain significant amounts of additional variance. The data did, however, show that cognitive predictors accounted for an additional 17% of the variance in OTDL-R performance after controlling for age, with processing speed being the only unique cognitive predictor of performance on the OTDL-R. Although the reason for this is not intuitively clear, as the OTDL-R is not a timed measure, psychomotor speed has emerged as a significant predictor of everyday functioning in prior studies (e.g., Tuokko et al., 2005). Contrary to our hypotheses, performance on the D-KEFS letter fluency was not a unique predictor of performance on any of the functional status measures.[5] While many prior studies have emphasized a link between executive functioning and IADL performance in older adult populations (e.g., Bell-McGinty et al., 2002; Cahn-Weiner, Boyle, & Malloy, 2002), the literature investigating cognitive predictors of functional status has yielded mixed results (Jefferson et al., 2006). In addition, the variance in functional status that can be attributed to cognitive correlates has varied widely across prior studies (e.g., 0% to 80%; Royall et al., 2007). While non-cognitive correlates controlled for in prior studies and neuropsychological variables used as cognitive predictors have differed widely, and may account for some of the varied findings, the current study highlights that variability in findings can also be attributed to choice of functional status measure.

Because the cognitive variables did not account for additional variance above age for the self-report IADL, EPT, and direct observation measures, this suggests that performance on the OTDL-R in this cognitively healthy aging population may be more closely tied with cognitive correlates than performances on the other functional status measures. This is further supported by the findings that while the cognitive predictors accounted for an additional 17% of the variance in OTDL-R performance, the functional status measures explained a non-significant 3% of the variance.

Of interest was the finding that both the self-report IADL and the EPT accounted for unique variance in the direct

---

(footnote continued)
the EPT, $B = -.52$, $t = -4.49$, $p < .001$. The functional status measures continued to account for a significant amount of the variance in the direct observation measure [$R^2 = .43$; $F(3,73) = 18.28$; $p < .001$], with both the EPT, $B = -.42$, $t = -4.49$, $p < .001$, and the self-report IADL, $B = -.30$, $t = -3.39$, $p = .001$, as well as the OTDL-R, $B = -.24$, $t = -2.63$, $p = .01$, serving as unique predictors. The functional status measures now also accounted for significant variance in the OTDL-R [$R^2 = .17$; $F(3,73) = 5.03$; $p = .003$], with the direct observation measure serving as a unique predictor, $B = -.35$, $t = -2.63$, $p = .01$.

---

[5] To examine for possible relationships with other executive functioning tests, we substituted the D-KEFS Design Fluency and the WAIS-III LNS subtest for the D-KEFS Letter Fluency as the cognitive predictor. The pattern of findings remained exactly identical. Furthermore, in no case did the D-KEFS Design Fluency subtest or the WAIS-III LNS subtest emerge as a unique predictor for a functional status measure.

observation measure, despite the fact that none of the cognitive variables predicted direct observation performance when controlling for age. Of further note, while age and cognitive variables accounted for 35% of the variance in the direct observation score, age and other functional status measures accounted for 47% of the variance. These findings are consistent with other studies which have shown poor links between neuropsychological testing scores and naturalistic action performance (e.g., Giovannetti, Libon, Buxbaum, & Schwartz, 2002). These findings further suggest that, in the cognitively healthy aging population, use of proxy functional status measures like a self-report IADL or the EPT, may provide more informed information regarding quality of everyday activity completion than neuropsychological testing data.

It has been suggested that cognitive and functional statuses reflect two different dimensions of performance, with individuals varying in both cognitive reserve and functional reserve (Loewenstein & Acevedo, 2010). The data indicate that there are other variables that need to be taken into account when assessing/predicting functional status as they may influence the relationship between capacity to complete a task and task implementation. For example, task context (e.g., presence of distractor items), social network, social skills (e.g., inappropriate behaviors), and use of compensatory strategies may significantly influence performances in the real-world environment and remain to be better investigated.

Regarding limitations, our sample of cognitively healthy older adults was predominantly Caucasian, highly educated and reported low rates of depressive symptomology. This contrast with other clinical and community-based samples in the literature and limits the generalizability of our results to other samples of older adults, including those suffering from cognitive or significant health problems. Findings from the regression analyses were also limited by study sample size, the limited battery of neuropsychological tests administered, and the specific neuropsychological measures chosen as the predictor variables to represent the cognitive constructs. In addition, while the data examined the relationship between traditional proxy measures of functional status, the functional domains assessed by each measure did not show complete overlap and this should be considered when interpreting the study findings. Furthermore, although participants in this study were in good physical condition, future studies should more carefully control for non-cognitive physical limitations that might influence everyday functioning and determine whether participants had successfully engaged in the tested activities at some point in their lives. Future studies may also want to include both self-report and informant-based evaluation of functional abilities, as lack of insight into functional performance has been found in some studies with normal aging populations (e.g., Suchy, Kraybill & Franchow, 2011). In addition, because variance in neuropsychological and functional status measures tends to be more constricted in cognitively healthy samples, these findings cannot be generalized to neurologic populations and future research is needed.

In summary, the study findings suggest that one must be cautious in making predictions about the quality of everyday activity completion in cognitively healthy older adults from specific cognitive functions. Our data suggest that within this cognitively healthy aging population, cognition may not be as strongly associated with everyday functioning as one might expect, or our neuropsychological tests may not be sensitive enough to show their true relationship. The data did, however, suggest that a self-report IADL and the performance-based EPT, both used as proxy measures for real-world functioning, may be useful measures for assessing everyday functional status in cognitively healthy older adults. Developing reliable and valid measures of functional status that can quickly and efficiently be given in the office and that accurately reflect everyday functioning continues to be an important goal for future research.

REFERENCES

Allaire, J.C., & Marsiske, M. (1999). Everyday cognition: Age and intellectual ability correlates. *Psychology and Aging*, *14*(4), 627–644.

Allaire, J.C., & Marsiske, M. (2002). Well- and ill-defined measures of everyday cognition: Relationship to older adults' intellectual ability and functional status. *Psychology and Aging*, *17*(1), 101–115.

Alexander, N.B., Guire, K.E., Thelen, D.G., Ashton-Miller, J.A., Schultz, A.B., Grunwalt, J.C., & Giordani, B. (2000). Self-reported walking ability predicts functional mobility performance in frail older adults. *Journal of the American Geriatrics Society*, *48*, 1408–1413.

Bell-McGinty, S., Podell, K., Franzen, M., Baird, A.D., & Williams, M.J. (2002). Standard measures of executive function in predicting instrumental activities of daily living in older adults. *International Journal of Geriatric Psychiatry*, *17*(9), 828–834.

Bertrand, R.M., & Willis, S.L. (1999). Everyday problem solving in Alzheimer's patients: A comparison of subjective and objective assessments. *Aging & Mental Health*, *3*, 281–293.

Brandt, J., & Folstein, M. (2003). *Telephone interview for cognitive status*. Lutz, FL: Psychological Assessment Resources, Inc.

Burton, C.A., Strauss, E., Bunce, D., Hunter, M.A., & Hultsch, D.F. (2009). Functional abilities in older adults with mild cognitive impairment. *Gerontology*, *55*, 570–581.

Cahn-Weiner, D.A., Boyle, P.A., & Malloy, P.F. (2002). Tests of executive function predict instrumental activities of daily living in community dwelling older individuals. *Applied Neuropsychology*, *9*(3), 187–191.

Cahn-Weiner, D.A., Farias, S.T., Julian, L., Harvey, D.J., Kramer, J.H., Reed, B.R., … Chui, H. (2007). Cognitive and neuroimaging predictors of instrumental activities of daily living. *Journal of the International Neuropsychological Society*, *13*(5), 747–757.

Cahn-Weiner, D., Malloy, P., Boyle, P., Marran, M., & Salloway, S. (2000). Prediction of functional status from neuropsychological tests in community-dwelling elderly individuals. *The Clinical Neuropsychologist*, *14*, 187–195.

Cho, C., Alessi, C., Cho, M., Aronow, H., Stuck, A., Rubenstein, L., & Beck, J. (1998). The association between chronic illness and functional change among participants in a comprehensive geriatric assessment program. *Journal of the American Geriatrics Society*, *46*, 677–682.

Cullum, C.M., Saine, K., Chan, L.D., Martin-Cook, K., Gray, K.F., & Weiner, M.F. (2001). Performance-based instrument to assess functional capacity in dementia: The Texas Functional Living Scale. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *14*, 103–108.

Dassel, K.B., & Schmitt, F.A. (2008). The impact of caregiver executive skills on reports of patient functioning. *The Gerontological Society of America*, *48*, 781–792.

Delis, D.C., Kaplan, E., & Kramer, J.H. (2001). *Delis-Kaplan Executive Function System: Examiner's manual*. San Antonio, TX: The Psychological Corporation.

Diehl, M., Marsiske, M., Horgas, A., Rosenberg, A., Saczynski, J., & Willis, S. (2005). The revised observed tasks of daily living: A performance-based assessment of everyday problem solving in older adults. *Journal of Applied Gerontology*, *24*, 211–230.

Farias, S.T., Harrell, E., Neumann, C., & Houtz, A. (2003). The relationship between neuropsychological performance and daily functioning in individuals with Alzheimer's disease: Ecological validity of neuropsychological tests. *Archives of Clinical Neuropsychology*, *18*, 655–672.

Farias, S.T., Mungas, D., & Jagust, W. (2005). Degree of discrepancy between self and other-reported everyday functioning by cognitive status: Dementia, mild cognitive impairment, and healthy elders. *International Journal of Geriatric Psychiatry*, *20*, 827–834.

Farias, S.T., Mungas, D., Reed, B., Harvey, D., Cahn-Weiner, D., & DeCarli, C. (2006). MCI is associated with deficits in everyday functioning. *Alzheimer Disease and Associated Disorders*, *20*, 217–223.

Galanos, A.N., Fillenbaum, G.G., Cohen, H.J., & Burchett, B.M. (1994). The comprehensive assessment of community dwelling elderly: Why functional status is not enough. *Aging Clinical and Experimental Research*, *6*, 343–352.

Giovannetti, T., Libon, D.J., Buxbaum, L.J., & Schwartz, M.F. (2002). Naturalistic action impairments in dementia. *Neuropsychologia*, *40*, 1220–1232.

Glosser, G., Gallo, J., Duda, N., de Vries, J.J., Clark, C.M., & Grossman, M. (2002). Visual perceptual functions predict instrumental activities of daily living in patients with dementia. *Neuropsychiatry, Neuropsychology, and Behavioral Neurology*, *15*, 198–206.

Goldstein, G. (1996). Functional considerations in neuropsychology. In R.J. Sbordone & C.J. Long (Eds.), *Ecological validity of neuropsychological testing* (pp. 75–89). Delray Beach, FL: GR Press/St Lucie Press.

Guccione, A., Felson, D., Anderson, J., Anthony, J., Zhang, Y., Wilson, P., … Kannel, W. (1994). The effects of specific medical conditions on the functional limitation of elders in the Framingham study. *American Journal of Public Health*, *84*, 351–358.

Hayes, T.L., Abendroth, F., Adamid, A., Pavela, M., Zitzelberger, T.A., & Kaye, J.A. (2008). Unobtrusive assessment of activity patterns associated with mild cognitive impairment. *Alzheimer's & Dementia*, *4*, 395–405.

Hughes, C.P., Berg, L., Danzinger, W.L., Coben, L.A., & Martin, R.L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry*, *140*, 566–572.

Inzarti, D., & Basile, A.M. (2003). Activities of daily living and global functioning. *International Psychogeriatrics*, *15*, 225–229.

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, E.G., & Petersen, R.C. (1996). Neuropsychological testing norms above age 55: COWAT, BNT, MAE TOKEN, WRAT-R Reading, AMNART, Stroop, TMT, and JLO. *The Clinical Neuropsychologist*, *10*, 262–278.

Jefferson, A.L., Barakat, L.P., Giovannetti, T., Paul, R.H., & Glosser, G. (2006). Object perception impairments predict instrumental activities of daily living dependence in Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology*, *28*(6), 884–897.

Jefferson, A.L., Byerly, L.K., Vanderhill, S., Lambe, S., Wong, S., Oznoff, A., & Karlawish, J.H. (2008). Characterization of activities of daily living in individuals with mild cognitive impairment. *The American Journal of Geriatric Psychiatry*, *16*(5), 375–383.

Kempen, G., Sterverink, N., Ormel, J., & Deeg, D. (1996). The assessment of ADL among frail elderly in an interview survey: Self-report versus performance based tests and determinants of discrepancies. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *51*, 254–260.

Lafortune, G., & Balestat, G. (2007). Trends in severe disability among elderly people: Assessing the evidence in 12 OECD countries and the future implications. *OECD Health Working Papers*, *No 26*, OECD Publishing.

Lawton, M., Moss, M., Fulcomer, M., & Kleban, M. (1982). A research and service oriented multilevel assessment instrument. *Journal of Gerontology*, *37*(1), 91–99.

Lewis, M.S., & Miller, L.S. (2007). Executive control functioning and functional ability in older adults. *The Clinical Neuropsychologist*, *21*(2), 274–285.

Loewenstein, D., & Acevedo, A. (2010). The relationship between instrumental activities of daily living and neuropsychological performance. In T.D. Marcotte & I. Grant (Eds.), *Neuropsychology of everyday functioning* (pp. 93–112). New York: The Guilford Press.

Loewenstein, D.A., Arguelles, S., Bravo, M., Freeman, R.Q., Arguelles, T., Acevedo, A., & Eisdorfer, C. (2001). Caregivers' judgments of the functional abilities of the Alzheimer's disease patient: A comparison of proxy reports and objective measures. *The Journals of Gerontology. Series B: Psychological Sciences and Social Sciences*, *56*, P78–P84.

Marcotte, T.D., Scott, J.C., Kamat, R., & Heaton, R.K. (2010). Neuropsychology and the prediction of everyday functioning. In T.D. Marcotte & I. Grant (Eds.), *Neuropsychology of everyday functioning* (pp. 5–38). New York: The Guilford Press.

Marson, D., & Hebert, K. (2006). Functional assessment. In D. Attix & K. Welsh-Bohmer (Eds.), *Geriatric Neuropsychology Assessment and Intervention* (pp. 158–189). New York: The Guilford Press.

McCue, M., Rogers, J.C., & Goldstein, G. (1990). Relationships between neuropsychological and functional assessment in elderly neuropsychiatric patients. *Rehabilitation Psychology*, *35*(2), 91–99.

Morris, J.C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, *43*, 2412–2414.

Myers, A., Holliday, P., Harvey, K., & Hutchinson, K. (1993). Functional performance measures: Are they superior to self-assessments? *Journal of Gerontology*, *48*, M196–M206.

Rashidi, P., Cook, D.J., Holder, L.B., & Schmitter-Edgecombe, M. (2011). Discovering activities to recognize and track in a smart environment. *IEEE Transactions on Knowledge and Data Engineering*, *23*, 527–539.

Reuben, D., Valle, L., Hays, R., & Siu, A. (1995). Measuring physical function in community dwelling older persons: A comparison of self-administered, interviewer-administered, and performance-based measures. *Journal of the American Geriatrics Society*, *43*, 17–23.

Richardson, E., Nadler, J., & Malloy, P. (1995). Neuropsychologic prediction of performance measures of daily living skills in geriatric patients. *Neuropsychology*, *9*, 565–572.

Royall, D.R., Cordes, J.A., & Polk, M. (1998). CLOX: An executive clock drawing task. *Journal of Neurology, Neurosurgery, and Psychiatry*, *64*, 588–594.

Royall, D.R., Lauterbach, E.C., Kaufer, D., Malloy, P., Coburn, K.L., & Black, K.J. (2007). The cognitive correlates of functional status: A review from the Committee on Research of the American Neuropsychiatric Association. *The Journal of Neuropsychiatry and Clinical Neurosciences*, *19*(3), 249–265.

Rueda, A., & Schmitter-Edgecombe, M. (2011). Everyday functioning in cognitively intact elderly compared to younger adults. Manuscript submitted for publication.

Smith, A. (1991). *Symbol Digit Modalities Test*. Los Angeles, CA: Western Psychological Services.

Suchy, Y., Kraybill, M., & Franchow, E. (2011). Instrumental activities of daily living among community-dwelling older adults: Discrepancies between self-report and performance are mediated by cognitive reserve. *Journal of Clinical and Experimental Neuropsychology*, *33*, 92–100.

Tuokko, H., Morris, C., & Ebert, P. (2005). Mild cognitive impairment and everyday functioning in older adults. *Neurocase*, *11*, 40–47.

Wechsler, D. (1997). *Wechsler Adult Intelligence Test-third edition*. New York, NY: The Psychological Corporation.

Williams, J.M. (1991). *Memory Assessment Scales professional manual*. Odessa: Psychological Assessment Resources, Inc.

Willis, S., & Marsiske, M. (1993). *Manual for the Everyday Problems Test*. University Park: Department of Human Development and Family Studies, Pennsylvania State University.

Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M.B., & Leirer, V.O. (1983). Development and validation of a geriatric depression rating scale: A preliminary report. *Journal of Psychiatric Research*, *17*, 37–49.

Zimmerman, S., & Magaziner, J. (1994). Methodological issues in measuring the functional status of cognitively impaired nursing home residents: The use of proxies and performance-based measures. *Alzheimer Disease and Associated Disorders*, *8*, S281–S290.

APPENDIX

A. Description of neuropsychological tests

*Neuropsychological Test*s

Telephone Interview for Cognitive Status: Participants complete simple tasks (e.g., count backwards from 20 to 1) designed to assess attention, orientation to place and time, memory and conceptualization.

Symbol Digits Modalities Test: Participants examine a series of nine geometric designs and for each symbol in the sequence, search a key for that symbol and substitute a number, either orally or in writing, for the symbol.

Memory Assessment Scale (MAS) List Learning subtest: For this standardized test assessing list acquisition abilities and memory, participants are read a 12-item list of semantically related words and asked to recall as many as possible. The list is read for up to 6 learning trials. After both a short-delay and a long-delay filled with other tasks, participants are asked to recall as many words as possible from the list.

Clox 1&2: Participants are asked to first draw (Clox 1) and then to copy (Clox 2) a clock, with the hands set to a specified time.

D-KEFS Letter Fluency subtest: Participants state aloud as many words as they can think of within a 60 second time period that begin with a specified letter (F, A, and S).

B. Description of activities and steps for accurate task completion

Description of Activity	Activity Completion Steps
Sweep and Dust: The participant sweeps the kitchen floor and dusts the dining and the living room using the supplies located in the kitchen closet.	1. Participant retrieves broom from supply closet 2. Participant retrieves duster from supply closet 3. Participant retrieves dust pan and brush from closet 4. Participant sweeps kitchen floor 5. Participant uses dust pan and brush 6. Participant dusts living room 7. Participant dusts dining room 8. Participant returns broom to supply closet 9. Participant returns duster to supply closet 10. Participants returns dust pan and brush to supply closet
Fill Medication Dispenser: The participant removes the items from kitchen cupboard and uses instructions taped to the kitchen cupboard door to fill the medication dispenser.	1. Participant retrieves materials from cupboard "A" 2. Participant reads instructions 3. Participant fills dispenser with medication 4. Participants returns supplies to kitchen cupboard
Write Birthday Card: The participant writes a birthday wish inside the birthday card and fills out a check in a suitable amount for a birthday gift, using the supplies located on the dining room table. He/she then places the card and the check in an envelope and appropriately addresses the envelope.	1. Participant writes birthday wish inside the card 2. Participant retrieves check 3. Participant writes check for birthday card 4. Participant puts birthday card and check in envelope 5. Participant addresses the envelope 6. Participant places envelope in the mail organizer 7. Participant returns other items to plastic bins
DVD Selection and Operation: The participant selects the DVD labeled "Good Morning America" located on the shelf below the TV and watches it on the TV. After watching it, the participant turns off the TV and returns the DVD to the shelf.	1. Participant selects correct DVD 2. Participant reads instructions next to the T.V. 3. Participant selects correct channel 4. Participants uses remote to turn on TV 5. Participant watches news clip 6. Participant turns off T.V. 7. Participant returns DVD to pile
Water Plants: The participant takes the watering can from the supply closet and lightly waters the 3 apartment plants, 2 of which are located on the kitchen windowsill and the third is located on the living room table. After finishing, he/she empties any extra water from the watering can into the sink and returns the watering can to the supply closet.	1. Participant retrieves watering can from supply closet 2. Participant fills watering can 3. Participant waters plants (windowsill) 4. Participant waters plants (coffee table) 5. Participant empties extra water into sink 6. Participant returns watering can to supply closet
Converse on Phone: The participant answers the phone when it rings and hangs up after finishing the conversation. The conversation includes several questions about the DVD clip that he/she watched as part of the "Watch DVD" activity.	1. Participant answers the phone 2. Participant answers questions over the phone 3. Participant hangs up/disconnects the phone
Prepare Meal: The participant uses the supplies located in the kitchen cupboard to prepare a cup of noodle soup according to the directions on the container. He/she also fills a glass with water using the pitcher of water located on the top shelf of the refrigerator.	1. Participant retrieves materials from cupboard "A" 2. Participant fills measuring cup with water 3. Participant boils water in microwave 4. Participant pours water into cup of noodles 5. Participant retrieves pitcher of water from refrigerator 6. Participant pours glass of water 7. Participant returns pitcher of water 8. Participant waits for water to simmer in cup of noodles 9. Participant brings all items to dining room table
Outfit Selection: The participant selects an outfit from the clothes closet to be worn on an important job interview. He/she then lays out the selected clothes on the living room couch.	1. Participant moves to hallway closet 2. Participant chooses correct outfit from closet 3. Participant lays out clothes on living room couch