Journal of the International Neuropsychological Society (2019), **25**, 530–545

doi:10.1017/S1355617719000122

Neuropsychological Test Norms in Cognitively Intact Oldest-Old

Zarui A. Melikyan,[1] Maria M. Corrada,[1,2,3] Malcolm B. Dick,[1] Christina Whittle,[1] Annlia Paganini-Hill,[2] AND Claudia H. Kawas[1,2,4]

[1]Institute for Memory Impairments and Neurological Disorders, University of California, Irvine, California
[2]Department of Neurology, University of California, Irvine, California
[3]Department of Epidemiology, University of California, Irvine, California
[4]Department of Neurobiology and Behavior, University of California, Irvine, California

(RECEIVED May 8, 2018; FINAL REVISION December 6, 2018; ACCEPTED January 7, 2019)

Abstract

Objectives: Individuals aged 90 or older (oldest-old), the fastest growing segment of the population, are at increased risk of developing cognitive impairment compared with younger old. Neuropsychological evaluation of the oldest-old is important yet challenging in part because of the scarcity of test norms for this group. We provide neuropsychological test norms for cognitively intact oldest-old. **Methods:** Test norms were derived from 403 cognitively intact participants of *The 90+ Study*, an ongoing study of aging and dementia in the oldest-old. Cognitive status of intact oldest-old was determined at baseline using cross-sectional approach. Individuals with cognitive impairment no dementia or dementia (according to DSM-IV criteria) were excluded. Participants ranged in age from 90 to 102 years (mean = 94). The neuropsychological battery included 11 tests (Mini-Mental Status Examination, Modified Mini-Mental State Examination, Boston Naming Test – Short Form, Letter Fluency Test, Animal Fluency Test, California Verbal Learning Test-II Short Form, Trail Making Tests A/B/C, Digit Span Forward and Backwards Test, Clock Drawing Test, CERAD Construction Subtests), and the Geriatric Depression Scale. **Results:** Data show significantly lower scores with increasing age on most tests. Education level, sex, and symptoms of depression were associated with performance on several tests after accounting for age. **Conclusions:** Provided test norms will help to distinguish cognitively intact oldest-old from those with cognitive impairment. (*JINS*, 2019, *25*, 530–545)

Keywords: Aged, 80 and over, Cognition, Neuropsychology, Reference values, Geriatric assessment

INTRODUCTION

The oldest-old (individuals aged 90 or older) are the fastest growing segment of the population. In the United States, the population of 90+ individuals is expected to triple by 2050, reaching 8.1 million people (United Nations Department of Economic and Social Affairs Population Division, 2017). Oldest-old individuals are at high risk of developing dementia (Corrada, Brookmeyer, Paganini-Hill, Berlau, & Kawas, 2010) and the ability to identify cognitive changes in this high-risk group is essential. However, distinguishing individuals with normal cognition from those with impaired cognition remains challenging because of the scarcity of appropriate test norms for this age group. Moreover, available test norms for cognitively normal oldest-old are limited by small sample sizes, small numbers of tests, or tests that are infrequently used by psychologists (Legdeur et al., 2017).

Correspondence and reprint requests to: Zarui A. Melikyan, 1511 Hewitt Hall, 843 Health Sciences Road, Irvine, CA 92697. E-mail: zmelikya@uci.edu

The present work addresses this gap by providing neuropsychological test norms that will help distinguish cognitively normal oldest-old from those with cognitive impairment [cognitive impairment with no dementia (CIND) and dementia]. Our earlier publication (Whittle et al., 2007) provided test norms to differentiate oldest-old without dementia (normal and CIND) from those with dementia. Inclusion of CIND participants in our previous normative publication resulted in lower means and larger variances of the normative values compared with norms derived from cognitively normal participants alone, and limited the ability to differentiate cognitively normal from mildly impaired individuals.

Here, we report test norms derived from one of the largest well-characterized cohorts of the oldest-old, *The 90+ Study*. Importantly, these new norms span a comprehensive battery of widely used cognitive tests (Rabin, Paolillo, & Barr, 2016). We developed norms by using a cross-sectional approach to determining cognitive status of the normative group, including individuals with normal cognition at

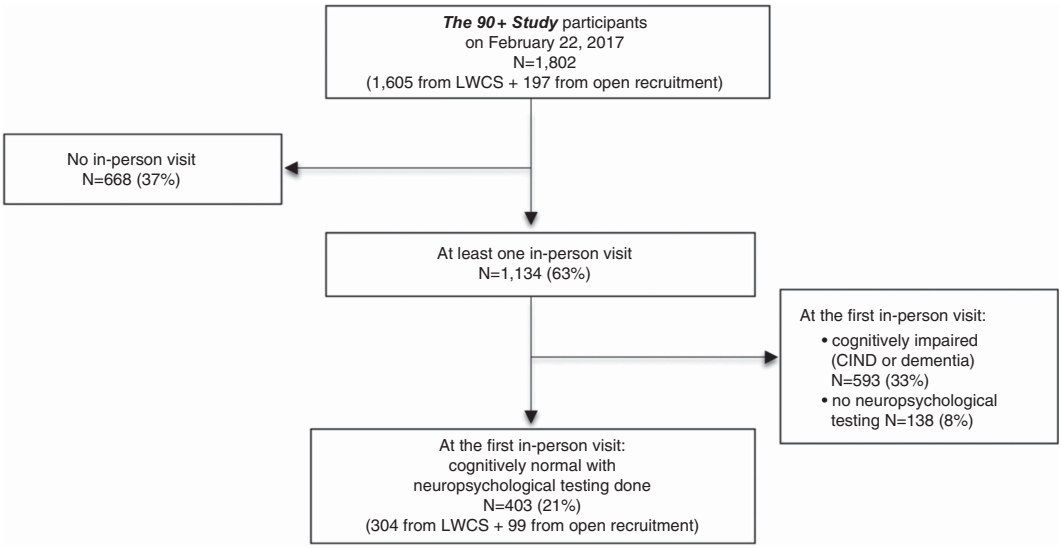

Fig. 1. Flowchart of participant inclusion. LWCS = Leisure World Cohort Study. All percentages are calculated from the total *90+ Study* cohort ($N = 1802$).

baseline, although they may have later developed cognitive impairment (Sliwinski, Lipton, Buschke, & Stewart, 1996). Using clinical diagnostic criteria we excluded individuals with CIND (Graham et al., 1997) and dementia (DSM-IV) (American Psychiatric Association, 1994) from the normative group.

METHODS

Study Procedures

We report results from a subset of participants of *The 90+ Study*, an ongoing longitudinal study of aging and dementia in people aged 90 or older. Participants of *The 90+ Study* were recruited from two groups: (1) survivors of the Leisure World Cohort Study (LWCS) (Paganini-Hill, Ross, & Henderson, 1986), a health survey study in the 1980s of the residents of Leisure World, a retirement community in Orange County, California, who were aged 90 or older on or after January 1, 2003, when enrollment into *The 90+ Study* commenced, and (2) 90+ residents of Orange County, California, who lived within a 2-hr drive of the study location, and joined the study through open recruitment (Melikyan et al., 2018).

Eligible individuals could participate in *The 90+ Study* at any of four levels: (1) in-person, (2) over the telephone, (3) through an informant, (4) LWCS participants who died before they themselves could participate in *The 90+ Study* were included if an informant provided information on medical, family history, and daily functioning. In-person participants undergo comprehensive semi-annual evaluations that include medical and family history, daily functioning, neurological examination, and neuropsychological testing. Based on participant's choice, visits are done at the study office or at home. We travel across the United States to test participants who have moved after enrollment.

The study was approved by the University of California Irvine's Institutional Review Board and all participants provided signed informed consent. Research was completed in accordance with the Helsinki Declaration.

Participants

Inclusion and exclusion criteria

This study reports on a subset of *The 90+ Study* participants who had at least one in-person evaluation and were determined by neurological examiners to have normal cognition at the first in-person evaluation. There were no other inclusion/exclusion criteria.

Of the 1,802 participants of *The 90+ Study* as of February 22, 2017 (Figure 1), 1134 (63%) had an in-person visit. Of these, 593 were classified as having CIND/dementia at the first in-person evaluation and an additional 138 had no neuropsychological testing done leaving 403 for analysis. These 403 individuals include 159 cognitively normal participants included in our previous publication (Whittle et al., 2007).

Data Collection Instruments

Background information and history

We collected information on demographics, medical history (participants were asked: "Have you ever been diagnosed with cardiovascular, cancer, psychiatric, neurological, or metabolic disorders?"), current mediations, living situation, and instrumental activities of daily living (IADL). Information on subjective cognitive decline was not collected.

Neuropsychological test battery

A neuropsychological test battery of 11 tests indexed language, word list memory, executive function, attention and

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

Table 1. Neuropsychological test battery in the order of administration

Tests in order of administration	Range of scores	Units
MMSE	0–30	Points
3MS	0–100	Points
Animal Fluency	0-max	No. of words in 1 min.
CVLT-II SF		
Trials 1–4	0–9	No. of words
Short Delay	0–9	No. of words
Clock Drawing	0–8	Points
Trail Making Test		
A	1–180, 0-max	Seconds, No. of errors
B	1–300, 0-max	Seconds, No. of errors
C	1–150, 0-max	Seconds, No. of errors
CVLT-II SF		
Long Delay	0–9	No. of words
Cued Long Delay	0–9	No. of words
Recognition	0–9	No. of words
CERAD Constructions	0–11	Points
BNT-Short	0–15	No. of items
Letter F Fluency	0-max	No. of words in 1 min.
Digit Span		
Forward	0–16	Points
Backward	0–14	Points
Geriatric Depression Scale	0–15	Points

Note. MMSE = Mini-Mental State Examination; 3MS = Modified Mini-Mental State Examination; CVLT-II SF = California Verbal Learning Test-II, Short Form; CERAD = The Consortium to Establish a Registry for Alzheimer's Disease; BNT-Short = Boston Naming Test, Short Form (15 items).

working memory, psychomotor speed, visual-spatial functions, construction; a questionnaire indexed symptoms of depression. The tests indexed different levels of cognitive ability while minimizing excessive floor and ceiling effects. Tests were administered in the order shown in Table 1 to maximize completion rates of the same tests in oldest-old participants who have high rates of incomplete testing due to fatigue. The average time to complete the entire battery was approximately 1 hour. Psychometrists, individuals with at least Bachelor's degree in psychology or related field and trained by a licensed neuropsychologist (M.B.D.), administered the tests in a standardized way.

Participants were asked to wear their eyeglasses and hearing aids during testing. In case of inability to complete a test due to sensory or motor impairment, a missing code indicated the reason for non-completion. Modifications, such as pairing printed and auditory stimuli and using enlarged boldface font for written information, were made to help compensate for sensory impairments. All test results, whether or not the whole battery was completed, were analyzed. Participants who did not complete the entire test battery were not excluded from analyses.

Cognitive screening tests included Modified Mini-Mental State Examination (3MS) (Teng & Chui, 1987) and Mini-Mental Status Examination (MMSE) (Folstein, Folstein, & McHugh, 1975). Most MMSE items are incorporated in the 3MS, and the addition of two items (which floor the participant is on and writing a sentence) to the 3MS made it possible to derive a total score for both tests. Two minor changes were made to the standard administration procedures: (1) the three to-be-remembered words were printed on three separate cards in 90-size font and shown one at a time while the examiner simultaneously repeated the words aloud, (2) the 60-s Animal Fluency test (Morris et al., 1989) was substituted for the 3MS 30-s task of naming four-legged animals.

Language was indexed using confrontational object naming, category (animals), and letter (F) (Gladsjo, Schuman, Miller, & Heaton, 1999; Heaton, Miller, Taylor, & Grant, 2004) fluencies. Object naming was indexed with the 15-item version of the Boston Naming Test (BNT-Short) (Fastenau, Denburg, & Mauer, 1998) to reduce administration time and fatigue. To avoid confusion with similar-sounding letters, a large "F" printed in 200-size font on a card was presented as a prompt.

Word list memory was indexed with California Verbal Learning Test - Second Edition, Short Form (CVLT-II SF) (Delis, Kramer, Kaplan, & Ober, 2000). Our modification was to present the words both verbally and visually (one at a time) during the four learning trials. A Short Delay Free Recall was administered following a 30-s interference task of counting backward from 100 by ones. After approximately 10 min of nonverbal tasks, the Long Delay Free Recall was administered and tests of cued-recall and yes/no recognition administered immediately thereafter.

Executive functioning and attention were indexed using the Trail Making Tests (TMT) Parts A and B using standard administration procedures (Reitan & Wolfson, 1993). Completion time limit was 180 s for TMT A and 300 s for TMT B. *Working memory* was indexed using Digit Span Forwards and Backwards from the Wechsler Adult Intelligence Scale-III (Wechsler, 1997). The administration and scoring followed standard procedures.

Psychomotor speed was indexed using a short and less tiring instrument developed by our group that is similar to the original Delis–Kaplan Executive Functioning System (D-KEFS) TMT Part C (Delis, Kaplan, & Kramer, 2001). Using the stimulus page from TMT Part A, we removed the numbers leaving the empty circles that we connected with a dotted line. We reversed the Part A starting and ending points, so that the Part A ending point (i.e., location of number 25) became the beginning position and the Part A starting point (i.e., location of the number 1) became the ending position. The participant's task was to trace over the dotted line, connecting the circles as quickly as possible using a marker. Completion time limit was 150 s.

Visual-spatial and constructional abilities were indexed using the Clock Drawing test and the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) Construction Test. In the Clock Drawing test, the participant was asked to

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

fill in a pre-drawn, 4-inch-diameter circle with numbers to represent a clock face and then to draw the hands at "ten after eleven." In the CERAD Construction Test the participant was asked to copy circle, four-sided diamond, intersecting rectangles, and cube.

Symptoms of depression were characterized using the Geriatric Depression Scale (GDS) (Yesavage et al., 1982-1983).

More detailed information on testing procedures and scoring is provided in Supplementary Materials.

Cognitive status assessment and diagnosis

Cognitive status was determined using: (1) a structured neurological examination; (2) the MMSE, 3MS, and Animal Fluency Test; (3) the Clinical Dementia Rating (CDR) scale (Morris, 1993); and (4) the Functional Activities Questionnaire (FAQ) (Pfeffer, Kurosaki, Harrah, Chance, & Filos, 1982). Participants were categorized based on the clinical diagnostic criteria as: (1) dementia, according to the Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) (American Psychiatric Association, 1994), that is, impaired performance on MMSE or 3MS subtests indexing at least two cognitive domains and inability to perform at least one IADL; (2) CIND, that is, impaired performance on MMSE or 3MS subtests or some difficulty in performing IADLs due to cognition, but not meeting criteria for a dementia diagnosis (Graham et al., 1997); or (3) normal cognition, that is, no substantial impairment on any cognitive domain (from subtests of MMSE, 3MS, or CDR) and no functional difficulties due to cognitive loss (from FAQ or CDR). Only individuals with normal cognition at their first in-person evaluation were included in the normative group.

Neurological examiners performed the cognitive status assessment and determined diagnostic classification at the end of the evaluation. No consensus diagnosis was used. Neurological examiners were physicians or nurse practitioners trained on the application of CIND and DSM-IV dementia diagnostic criteria by a licensed geriatric neurologist (C.K.).

We report norms on MMSE, 3MS, and animal fluency, that were used in determination of cognitive status, for two reasons: (1) these test scores were not the only criterion for cognitive diagnosis, another being performance in IADLs; (2) these tests are frequently used in aging and dementia settings and have low non-completion rates, making their norms useful.

If 4 or fewer scores on MMSE or 12 or fewer scores on 3MS were missing due to sensory, motor, or other difficulties, proportional scores were computed: proportional MMSE score = ((30*MMSE total)/(30-MMSE number of missing points)), proportional 3MS score = ((100*3MS total)/(100-3MS number of missing points)). This calculation assumes that the score obtained without completing all items would be proportionally equal to the score that would have been obtained if all items had been completed. The fewer scores missing, the more accurately the proportional score represents the theoretical total score; therefore, cutoffs for the number of missing items were established. If more than 4 scores in the MMSE or more than 12 scores in the 3MS were missing, proportional scores were not computed.

Data Analysis

Means, standard deviations, and percentiles (5, 10, 25, 50, 75, 90, and 95 percentiles) are reported for each test. For ease of use and comparison, norms are provided for the same age groups as in our previous report (Whittle et al., 2007): 90–91, 92–94, and ≥95 years. The effect of age was assessed by regression analysis with age as a continuous variable. The age-adjusted independent effects of sex, education [the same categories as in our previous work (Whittle et al., 2007): high school or less, some college to college graduate, at least some graduate school were used for consistency and ease of comparison], and GDS score (<4 *vs.* ≥4) were assessed by multivariable regression analyses. Effect sizes are reported using Cohen's *d* (Cohen, 1988). To compare characteristics among the age groups, we used Fisher's exact tests for categorical variables and *t* tests and analyses of variance for continuous variables. SAS version 9.4 (SAS Institute Inc., Cary, NC) was used for statistical analyses.

The study provides age norms, norms by sex and education for the tests with significant sex/education effects after adjusting for age, and missing data. Norms for men and women separately and optional scores (performance on subtests and training samples, cued responses, and errors) are provided in the Supplementary Material (Supplementary Tables 1–7).

RESULTS

Group Demographics and Health

The sample of 403 cognitively normal participants (283 women and 120 men) has an average age of 94 years (range, 90–102 years) (Table 2). Most participants were Caucasian (98.5%), well-educated (78% were educated beyond high school), and lived by themselves (63%). Education did not differ significantly among the three age groups (90–91, 92–94, ≥95) ($p = .79$, Fisher's exact test).

The most frequent health problems were history of hypertension (62%), heart disease (49%), and non-skin cancer (33%), with no significant differences in prevalence among the three age groups ($p = .52$, Fisher's exact test). Although 9% of participants reported receiving a diagnosis of depression, over 20% had an elevated depression score (GDS ≥ 4). The proportion of participants with GDS ≥ 4 increased significantly with age ($F(1,233) = 5.68$; $p = .02$). Reporting a diagnosis of depression also increased with age, although not significantly ($p = .11$, Fisher's exact test).

At the time of testing 83 (21%) participants reported taking psychoactive medications (narcotic analgesics, general anesthetics, anxiolytics, sedatives, hypnotics, CNS stimulants, antidepressants, antipsychotics, antiparkinsonian) or anti-dementia medications (cholinesterase inhibitors or

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

Table 2. Characteristics of study participants

Characteristic	Entire sample No. (%)	90–91 years No. (%)	92–94 years No. (%)	≥ 95 years No. (%)
Sex				
Women	283 (70.2)	85 (69.1)	118 (70.7)	80 (70.8)
Men	120 (29.8)	38 (30.9)	49 (29.3)	33 (29.2)
Race				
Caucasian	397 (98.5)	121 (98.4)	164 (98.2)	112 (99.1)
Asian/Pacific Islander	3 (.7)	1 (.8)	1 (.6)	1 (0.9)
Spanish/Hispanic/Latino	3 (.7)	1 (.8)	2 (1.2)	0
Age (years)	403 (100)	123 (30.5)	167 (41.4)	113 (28.1)
Residence				
At home alone	253 (62.8)	84 (68.3)	104 (62.3)	65 (57.5)
At home with spouse/relatives/friends	103 (25.6)	33 (26.8)	43 (25.7)	27 (23.9)
Institution/group home/at home with paid caregiver	47 (11.7)	6 (4.9)	20 (12.0)	21 (18.6)
Education				
High school graduate or less	90 (22.3)	27 (22.0)	41 (24.6)	22 (19.5)
Some college to college graduate	195 (48.4)	62 (50.4)	74 (44.9)	58 (51.3)
Some graduate school to graduate/professional degree	118 (29.3)	34 (27.6)	51 (30.5)	33 (29.2)
Geriatric Depression Scale				
< 4 depressive symptoms	268 (79.1)	87 (70.7)	115 (68.9)	66 (58.4)
≥ 4 depressive symptoms	71 (20.9)	17 (13.8)	26 (15.6)	28 (24.8)
Medical history				
Hypertension	248 (62.3)	75 (61.0)	104 (62.3)	69 (61.1)
Heart disease[a]	197 (48.9)	62 (50.4)	77 (46.1)	58 (51.3)
Cancer (other than skin)	131 (32.5)	40 (32.5)	57 (34.1)	34 (30.1)
Depression	35 (8.8)	7 (5.7)	13 (7.8)	15 (13.3)
Stroke	27 (6.7)	8 (6.5)	11 (6.6)	8 (7.1)
Diabetes	24 (6.0)	9 (7.3)	9 (5.4)	6 (5.3)
Psychoactive medications				
All psychoactive medications[b]	83 (20.6)	25 (20.3)	33 (19.8)	25 (22.1)
Anti-dementia medications[c]	6 (1.5)	2 (1.6)	2 (1.2)	2 (1.8)

Note. All percentages are column percentages out of total sample of 403 participants.
[a]Heart disease includes: coronary artery disease, myocardial infarction, atrial fibrillation or other arrhythmias, heart valve disease, and congestive heart failure.
[b]All psychoactive medications include: narcotic analgesics, general anesthetics, anxiolytics, sedatives, hypnotics, CNS stimulants, antidepressants, anti-psychotics, anti-parkinsonian, and anti-dementia medications.
[c]Anti-dementia medications include: cholinesterase inhibitors and/ or NMDA antagonist.

NMDA antagonist). Use of psychoactive medication was not significantly different among the three age groups ($p = .88$, Fisher's exact test). Only 6 (1.5%) participants were taking anti-dementia medications with no difference among the three age groups ($p = .11$, Fisher's exact test).

Effects of Age and Age-Adjusted Effects of Education, Sex, and Depressive Symptoms on Test Scores

With increasing age, total scores on MMSE, 3MS, BNT-Short number of spontaneous correct responses (henceforth listed as BNT-Short for brevity), animal fluency, free recall trials (including short and long delays) in CVLT-II SF, TMT A, and clock drawing test were significantly lower (Table 3).

After adjusting for age, individuals with more education scored significantly higher than those with less education on MMSE, 3MS, BNT-Short, animal and letter F fluencies, and CERAD Construction (Table 4).

After adjusting for age, men scored higher than women on BNT-Short, whereas women scored significantly higher than men on the MMSE and CVLT-II SF (Trials 2, 3, 4, Sum of Trials 1–4, short- and long-delay free recall). Effect sizes as measured by Cohen's *d* were small to medium (.25 to .36) (Table 5).

A higher GDS score was significantly associated with lower scores on 3MS, BNT-Short, animal and letter F fluencies, CVLT-II SF Trial 4, short- and long-delay free recall, and TMT A (results not shown).

Adjustment for education did not alter the effects of sex and GDS score on test scores.

Comparison of Participants Who Did and Did Not Complete All the Tests

Not all participants completed all tests, primarily due to fatigue, sensory impairments, or time constraints (Table 6). Hearing problems accounted for non-completion in 0.8–3.5% of participants (depending on the test), but the

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

Table 3. Raw neuropsychological test scores (mean, *SD*, percentiles) by age group

Test	Age group[a]	No.[b]	Mean	*SD*	5%	10%	25%	50%	75%	90%	95%	*B* ± *SE* /*t* /*p*[c]
MMSE	90–91	123	28.0	1.7	25	26	27	28	29	30	30	−.11 ± .03 / −3.31 / < .01
Total score	92–94	166	28.2	1.7	25	26	27	28	30	30	30	
	≥ 95	113	27.4	1.9	24	25	26	28	29	30	30	
	Overall	402	27.9	1.8	25	26	27	28	29	30	30	
3MS												
Total score	90–91	120	94.6	4.1	86	89	92	96	98	99	100	− .28 ± .09 / − 3.19 / < .01
	92–94	155	94.4	4.4	85	89	92	96	98	99	100	
	≥ 95	102	93.1	4.9	84	88	91	94	96	98	99	
	Overall	377	94.1	4.5	85	88	92	95	97	99	100	
BNT-Short												
Total correct												
	90–91	97	13.2	1.8	10	11	12	14	14	15	15	− .19 ± .04 / − 4.24 / < .01
	92–94	131	12.8	1.9	9	10	11	13	14	15	15	
	≥ 95	76	12.1	2.0	9	9	11	12	14	15	15	
	Overall	304	12.8	1.9	9	10	11	13	14	15	15	
Animal Fluency												
Total correct												
	90–91	122	14.8	4.2	9	10	12	14	17	20	23	− .19 ± .08 / − 2.25 /.02
	92–94	166	14.7	3.8	9	10	12	14	18	20	22	
	≥ 95	112	13.8	4.0	8	9	11	13	16	19	22	
	Overall	400	14.5	4.0	9	10	12	14	17	20	22	
Letter F Fluency												
Total correct												
	90–91	104	12.9	4.2	6	8	10	13	15	18	20	.02 ± .09 /.27 /.79
	92–94	138	12.7	3.9	7	7	10	13	16	17	19	
	≥ 95	93	12.7	4.5	6	7	10	12	15	19	22	
	Overall	335	12.8	4.2	7	8	10	13	15	18	20	
CVLT-II SF												
Trial 1	90–91	107	4.9	1.5	2	3	4	5	6	7	7	− .08 ± .03 / − 2.65 /.01
Number of words	92–94	150	4.9	1.6	2	3	4	5	6	7	7	
	≥ 95	96	4.4	1.3	2	3	4	4	5	6	7	
	Overall	353	4.8	1.5	2	3	4	5	6	7	7	
Trial 4	90–91	106	7.8	1.2	5	6	7	8	9	9	9	− .09 ± .03 / − 3.30 / < .01
Number of words	92–94	149	7.7	1.2	5	6	7	8	9	9	9	
	≥ 95	96	7.2	1.3	5	5	6	8	8	9	9	
	Overall	351	7.6	1.3	5	6	7	8	9	9	9	
Sum Trials1-4	90–91	106	27.0	4.5	19	20	24	27	30	32	34	− .30 ± .09 / − 3.12 / < .01
Number of words	92–94	149	26.9	4.4	18	20	25	27	30	32	33	
	≥ 95	96	24.8	4.4	17	19	21	25	29	31	32	
	Overall	351	26.3	4.5	18	20	23	27	30	32	33	
Short Delay Recall	90–91	106	7.4	1.4	5	5	7	8	9	9	9	− .09 ± .03 / − 3.01 /.003
Number of words	92–94	149	7.3	1.4	5	6	6	8	8	9	9	
	≥ 95	96	6.8	1.5	4	5	6	7	8	9	9	
	Overall	351	7.2	1.5	4	5	6	7	8	9	9	
Long Delay Recall	90–91	105	6.9	1.9	3	4	6	7	8	9	9	− .15 ± .04 / − 3.62 / < .01
Number of words	92–94	148	6.7	1.8	3	4	6	7	8	9	9	
	≥ 95	96	5.8	2.1	1	3	5	6	8	8	9	
	Overall	349	6.5	2.0	3	4	5	7	8	9	9	

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

Table 3. (*Continued*)

Test	Age group[a]	No.[b]	Mean	*SD*	5%	10%	25%	50%	75%	90%	95%	*B* ± *SE* /*t* /*p*[c]
Trail Making Test A												
Seconds	90–91	102	57.2	25.1	99	87	71	49	39	35	31	1.09 ± .55 /1.99 /.05
	92–94	130	55.3	18.8	88	82	66	51	42	34	30	
	≥ 95	78	63.4	29.0	128	103	74	57	45	38	33	
	Overall	310	58.0	24.0	104	87	69	53	42	35	31	
Trail Making Test B												
Seconds	90–91	87	139.7	57.6	244	216	173	133	94	68	64	2.21 ± 1.50 /1.47 /.14
	92–94	109	141.1	55.5	243	220	169	127	103	83	75	
	≥ 95	59	151.9	61.0	261	235	177	143	111	83	74	
	Overall	255	143.2	57.5	250	224	173	132	103	78	68	
Trail Making Test C												
Seconds	90–91	99	24.5	9.9	45.0	37.0	31.0	22.0	18.0	14.0	13.0	.46 ± .29 /1.60 /.11
	92–94	124	25.8	12.7	50.0	42.0	31.5	22.0	18.0	13.0	12.0	
	≥ 95	70	27.0	12.9	55.0	44.5	32.0	24.0	18.0	14.5	13.0	
	Overall	293	25.7	11.9	48.0	40.0	31.0	23.0	18.0	14.0	13.0	
Digit Span Forward + Backwards												
Total score	90–91	86	15.2	3.3	10	12	13	15	17	20	22	.06 ± .07 /.85 /.39
	92–94	103	14.7	2.8	11	12	13	14	16	18	20	
	≥ 95	65	15.4	2.7	11	11	14	16	17	19	20	
	Overall	254	15.0	3.0	11	12	13	15	16	19	21	
Clock Drawing												
Total score	90–91	110	6.2	1.7	4	4	5	7	8	8	8	− .12 ± .04 / − 3.13 / < .01
	92–94	141	5.9	1.7	3	3	5	6	7	8	8	
	≥ 95	90	5.3	1.9	2	3	4	5	7	8	8	
	Overall	341	5.8	1.8	3	3	4	6	7	8	8	
CERAD Construction Total score												
	90–91	99	9.0	1.4	7	7	8	9	10	11	11	− .01 ± .03 / − .42 /.68
	92–94	125	9.4	1.2	7	8	8	10	10	11	11	
	≥ 95	81	9.1	1.1	8	8	8	9	10	11	11	
	Overall	305	9.2	1.3	7	8	8	9	10	11	11	

Note. MMSE = Mini-Mental State Examination; 3MS = Modified Mini-Mental State Examination; BNT-Short = Boston Naming Test, Short Form (15 items); CVLT-II SF = California Verbal Learning Test-II, Short Form; CERAD = The Consortium to Establish a Registry for Alzheimer's Disease. MMSE and 3MS were used in determination of cognitive status.

[a]In years.

[b]Number of participants does not always total 403 as not all the participants completed all the tests.

[c]*B* ± *SE* /*t* /*p* = parameter estimate ± standard error /*t*-value /*p*-value from linear regression analysis with age as a continuous variable.

non-completion rate did not differ among the three age groups ($p = .33$, Fisher's exact test) (Table 7). Non-completion due to motor symptoms (such as tremor) significantly increased with age from 0% in the 90–91 group to 1–2.4% in the two older age groups ($p < .01$, Fisher's exact test). Vision impairment accounted for approximately 6% of non-completion in the two younger groups and significantly increased to approximately 16% in the ≥95 age group ($p = .01$, Fisher's exact test).

On the cognitive screening tests, 363 participants (90%) completed all MMSE items and 39 participants (10%) had 1–4 missing scores. The average MMSE score for participants who completed all items (mean = 27.9; $SD = 1.7$) did not differ from the proportional MMSE score computed for participants with 1–4 missing scores (mean = 27.4; $SD = 2.2$; $t(43) = 1.54$; $p = .13$). All 3MS items were completed by 362 participants (96%); 15 participants (4%) had 1–12 missing scores. The average 3MS score for participants who completed all 3MS items (mean = 94.2; $SD = 4.4$) was higher than the proportional 3MS score computed for participants with 1–12 missing scores (mean = 91.3; $SD = 5.6$; $t(375) = 2.49$; $p = .01$).

Within the entire testing battery, completion rates were high for tests administered first: MMSE (>99%), 3MS (94%), and Animal Fluency (99%). In comparison, tests administered toward the end of the battery were least likely to be completed: TMT B (63%) and Digit Span Test (63%). MMSE and 3MS scores were significantly higher among

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

Table 4. Raw neuropsychological test scores (mean, *SD*) by age group for different education levels

		Education Level			
Test	Age group	≤ High school (HS) Mean (*SD*)	≤ College Mean (*SD*)	> College Mean (*SD*)	$B \pm SE$ /*t* /*p*[a]
MMSE		*N* = 90	*N* = 194	*N* = 118	.29 ± .12 /2.39 /.02
Total score	90–91	27.5 (2.1)	28.0 (1.6)	28.3 (1.3)	
	92–94	27.8 (1.9)	28.2 (1.7)	28.3 (1.7)	
	≥ 95	26.9 (2.1)	27.6 (1.7)	27.5 (2.2)	
	Overall	27.5 (2.0)	28.0 (1.7)	28.1 (1.8)	
3MS		*N* = 85	*N* = 185	*N* = 109	1.45 ± .31 /4.62 / < .01
Total score	90–91	92.5 (4.3)	94.7 (4.3)	96.1 (3.0)	
	92–94	92.5 (5.2)	94.6 (3.9)	95.8 (4.0)	
	≥ 95	91.7 (7.2)	93.4 (4.2)	93.3 (4.3)	
	Overall	92.3 (5.4)	94.3 (4.2)	95.2 (4.0)	
BNT-Short		*N* = 61	*N* = 157	*N* = 86	.44 ± .16 /2.80 / < .01
Total correct					
	90–91	12.3 (2.7)	13.6 (1.4)	13.1 (1.5)	
	92–94	12.3 (2.0)	12.8 (2.0)	13.3 (1.6)	
	≥ 95	11.2 (2.1)	12.3 (1.8)	12.2 (2.0)	
	Overall	12.1 (2.3)	12.9 (1.9)	13.0 (1.8)	
Animal Fluency		*N* = 90	*N* = 193	*N* = 117	.79 ± .28 /2.86 / < .01
Total correct					
	90–91	13.3 (4.0)	14.5 (4.1)	16.6 (4.2)	
	92–94	14.2 (3.8)	14.6 (3.7)	15.3 (4.0)	
	≥ 95	13.8 (3.5)	13.7 (3.3)	14.1 (5.3)	
	Overall	13.8 (3.8)	14.3 (3.7)	15.4 (4.5)	
Letter F Fluency		*N* = 68	*N* = 168	*N* = 99	.99 ± .32 /3.10 / < .01
Total correct					
	90–91	11.1 (3.1)	13.4 (4.4)	13.1 (4.4)	
	92–94	11.1 (3.8)	13.1 (4.0)	13.2 (3.5)	
	≥ 95	11.1 (4.5)	12.8 (4.5)	13.6 (4.3)	
	Overall	11.1 (3.8)	13.1 (4.3)	13.3 (3.9)	
CERAD Construction		*N* = 90	*N* = 195	*N* = 118	.22 ± .10 /2.17 /.03
Total score					
	90–91	8.8 (1.4)	9.4 (1.3)	8.8 (0.8)	
	92–94	8.9 (1.4)	9.2 (1.3)	9.2 (1.2)	
	≥ 95	9.5 (1.1)	9.6 (1.2)	9.2 (1.1)	
	Overall	9.1 (1.2)	9.1 (1.3)	9.5 (1.1)	

Note. Scores are provided only for tests for which education level significantly contributed to test performance after controlling for age. MMSE = Mini-Mental State Examination; 3MS = Modified Mini-Mental State Examination; BNT-Short = Boston Naming Test, Short Form (15 items).
[a]$B \pm SE$ /*t* /*p* = parameter estimate ± standard error /*t*-value /*p*-value from linear regression analysis with age as continuous and education as categorical variable (high school or less, some college to completed college, some graduate school to completed graduate school).

those who completed compared with those who did not complete select neuropsychological tests (BNT-Short, CVLT, TMT B and C, Digit Span, CERAD for MMSE and 3MS; TMT A also for 3MS) (Table 8).

DISCUSSION

This report extends the available neuropsychological test norms for cognitively normal individuals aged 90 or older. We report norms by age group, sex, and education, which, along with symptoms of depression, influence test performance. These norms allow differentiation of cognitively normal individuals from those with cognitive impairment (CIND or dementia). In contrast, the norms in our earlier publication (Whittle et al., 2007) were helpful in distinguishing between oldest-old with dementia and those without dementia (cognitively normal and CIND).

Consistent with our previous publication (Whittle et al., 2007) and other reports (Dore, Elias, Robbins, Elias, &

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

Table 5. Raw neuropsychological test scores (mean, *SD*) by age group for women and men

Test	Age group	Sex: Women Mean (*SD*)	Sex: Men Mean (*SD*)	B ± SE /t /*p*[a]	*Effect size Cohen's d*
MMSE		*N* = 282	*N* = 120	.59 ± .19 /3.11 / < .01	–.33
Total score	90–91	28.0 (1.7)	28.0 (1.7)		
	92–94	28.5 (1.5)	27.4 (2.0)		
	≥ 95	27.6 (1.9)	27.0 (2.0)		
	Overall	28.1 (1.7)	27.5 (1.9)		
BNT-Short		*N* = 212	*N* = 103	–.65 ± .23 /–2.76 / < .01	.34
Total correct	90–91	13.0 (2.0)	13.5 (1.3)		
	92–94	12.7 (2.0)	13.2 (1.8)		
	≥ 95	11.8 (2.0)	12.8 (1.7)		
	Overall	12.6 (2.0)	13.2 (1.6)		
CVLT-II SF Trial 2		*N* = 283	*N* = 120	.36 ± .16 /2.23 /.03	–.25
Number of words	90–91	7.0 (1.4)	6.3 (1.4)		
	92–94	6.9 (1.3)	6.6 (1.2)		
	≥ 95	6.2 (1.5)	6.3 (1.3)		
	Overall	6.8 (1.4)	6.4 (1.3)		
CVLT-II SF Trial 3		*N* = 283	*N* = 120	.46 ± .15 /3.04 / < .01	–.35
Number of words	90–91	7.6 (1.3)	7.1 (1.3)		
	92–94	7.6 (1.1)	7.0 (1.4)		
	≥ 95	7.1 (1.4)	6.8 (1.3)		
	Overall	7.4 (1.3)	7.0 (1.4)		
CVLT-II SF Trial 4		*N* = 283	*N* = 120	.45 ± .14 /3.14 / < .01	–.36
Number of words	90–91	8.0 (1.1)	7.5 (1.1)		
	92–94	7.9 (1.2)	7.4 (1.2)		
	≥ 95	7.3 (1.2)	7.0 (1.5)		
	Overall	7.7 (1.2)	7.3 (1.3)		
CVLT-II SF Trials 1-4		*N* = 248	*N* = 103	1.53 ± .52 /2.97 / < .01	–.33
Number of words	90–91	27.6 (4.4)	25.5 (4.5)		
	92–94	27.3 (4.3)	25.7 (4.4)		
	≥ 95	25.1 (4.6)	24.2 (3.9)		
	Overall	26.8 (4.5)	25.3 (4.3)		
CVLT-II SF Short Delay		*N* = 248	*N* = 103	.55 ± .17 /3.34 / < .01	–.38
Number of words	90–91	7.7 (1.4)	6.9 (1.5)		
	92–94	7.5 (1.4)	6.9 (1.2)		
	≥ 95	6.8 (1.4)	6.6 (1.9)		
	Overall	7.4 (1.4)	6.8 (1.5)		
CVLT-II SF Long Delay		*N* = 246	*N* = 103	.80 ± .22 /3.55 / < .01	–.40
Number of words	90–91	7.4 (1.7)	5.9 (2.0)		
	92–94	6.9 (1.8)	6.1 (1.8)		
	≥ 95	5.7 (2.1)	5.8 (2.3)		
	Overall	6.7 (1.9)	5.9 (2.0)		

Note. Scores provided only for tests for which gender significantly contributed to test performance after controlling for age. MMSE = Mini-Mental State Examination; BNT-Short = Boston Naming Test, Short Form; CVLT-II SF = California Verbal Learning Test-II, Short Form.
[a]*B* ± *SE* /*t* /*p* – parameter estimate ± standard error /t-value /*p*-value from linear regression analysis with age as continuous and sex as categorical variable.

Brennan, 2007; Elias, Elias, D'Agostino, Silbershatz, & Wolf, 1997; Harada, Natelson Love, & Triebel, 2013), the current analysis shows that performance on screening measures and on tests indexing attention, language, verbal memory, and construction declines significantly with advancing age. Age-related change in cognitive performance including decline in speeded aspects of activity (Eckert, Keren, Roberts, Calhoun, & Harris, 2010), failure to suppress

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

Table 6. Percent of participants[a] who completed and did not complete each neuropsychological test and reasons for non-completion

		Reasons for non-completion						
Cognitive tests and domains they index	Completed	Fatigue	Hearing	Vision	Cognition[b]	Refused	Out of Time[c]	Other[d]
Cognitive Screening Tests								
MMSE Total score	99.8	0.0	0.0	0.0	0.0	0.0	0.0	0.3
3MS Total score	93.6	0.0	0.0	0.0	0.0	0.0	0.0	6.5
Language								
BNT-Short Total correct	75.4	4.5	0.0	7.9	0.0	1.0	10.2	0.3
Animal Fluency Total correct	99.3	0.3	0.0	0.0	0.0	0.0	0.3	0.3
Letter F Fluency Total correct	83.1	5.5	0.0	0.0	0.0	1.5	9.4	0.5
Verbal Memory								
CVLT-II SF Number of words	87.6	2.5	0.5	0.0	0.0	0.7	8.7	0.0
Executive Function/Attention								
Trail Making Test A Seconds	76.9	3.2	0.0	9.2	0.0	1.5	7.4	1.7
Trail Making Test B Seconds	63.3	3.2	0.0	9.7	7.4	2.0	9.9	4.5
Psychomotor Speed								
Trail Making Test C Seconds	72.7	3.7	0.0	8.9	0.0	3.0	10.4	1.2
Working Memory								
Digit Span Test Total score	63.0	5.7	2.2	0.0	0.0	1.0	27.5	0.5
Construction								
Clock Drawing Total score	84.6	2.5	0.0	6.7	0.0	1.0	4.2	1.0
CERAD Construction Total score	75.7	4.0	0.0	7.7	0.0	1.7	9.4	1.5

Note. MMSE = Mini-Mental State Examination; 3MS = Modified Mini-Mental State Examination; BNT-Short = Boston Naming Test, Short Form (15 items); CVLT-II SF = California Verbal Learning Test-II, Short Form; CERAD = The Consortium to Establish a Registry for Alzheimer's Disease.
[a]Percent of participants out of the total 403 participants in the study sample.
[b]Could not understand instructions, became confused, forgot instructions.
[c]Ran out of time for the entire neuropsychological assessment, not individual test.
[d]Equipment error, tester error, other physical impairment of participant, e.g., tremor, alternate test given, quit after starting.

irrelevant information (Dumas & Hartman, 2008), and decreased use of strategies to improve learning and memory (Davis et al., 2013) is thought to be associated with structural and functional brain changes in older adults (Hafkemeijer et al., 2014; Liu et al., 2017). Counter to our previous report (Whittle et al., 2007), the current sample showed no age effect on TMT B or Digit Span Backwards. In *The 90+ Study* group and others (Rasmusson, Zonderman, Kawas, & Resnick, 1998) CIND explains larger proportion of variance in test performance than age.

In the current sample, education, sex, and symptoms of depression contributed to test performance independently of age. Similarly to others (Au et al., 2004; Dore et al., 2007; Elias et al., 1997; Ganguli et al., 2010; Saykin et al., 1995), we found an effect of education on cognitive screening tests and on tests that index naming, verbal fluency and construction. As education is implicated in cognitive reserve, slower age-related cognitive decline, and overall test-wiseness (de Azeredo Passos et al., 2015; Gasquoine, 2009; Stern, 2012), it can contribute to test performance.

In the current group, men scored significantly higher than women on the test indexing naming, but lower on the cognitive screening tests and verbal memory. Higher scores on the naming test in men than women have been reported previously, with no consensus on the mechanisms of these differences. Factors that have been explored include IQ and white matter changes (Hall, Vo, Johnson, Wiechmann, & O'Bryant, 2012). Although men in our group were slightly more educated than women, education did not explain sex differences in test performance. Higher performance of women than men on cognitive screening tests and tests indexing verbal memory has been demonstrated previously and ascribed to different approaches to encoding and learning in men and women or hormonal factors (Gale, Baxter, Connor, Herring, & Comer, 2007; Hogervorst, Rahardjo, Jolles, Brayne, & Henderson, 2012; Rosselli, Tappen, Williams, & Salvatierra, 2006). Although the observed effect sizes of sex differences in test performance were not large, use of sex-specific norms is recommended when available.

The well-documented association of elevated scores on depression measures with lower cognitive performance (Koenig, Bhalla, & Butters, 2014; Morimoto & Alexopoulos, 2013) was observed in our group on cognitive screening tests and on tests that index memory, verbal fluency, and attention. This could be related to poor effort, underlying subclinical dementia, or disruption in structural and functional brain integrity due to factors such as cerebrovascular pathology (Weisenbach, Boore, & Kales, 2012).

The prevalence of self-reported health problems in our group is similar to other reports for the oldest-old (Lee, Go, Lindquist, Bertenthal, & Covinsky, 2008; Nosraty, Sarkeala, Hervonen, & Jylhä, 2012). We found no differences among

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

Table 7. Percent of participants who did not complete tests due to sensory or motor impairment by age group[a]

Cognitive test	Age group	Reasons for non-completion: Hearing (%)	Vision (%)	Physical impairment (%)
BNT-Short Total correct	90–91	—	5.7	—
	92–94	—	5.4	—
	≥ 95	—	14.2	—
Trail Making Test A Seconds	90–91	—	5.7	0.0
	92–94	—	6.6	2.4
	≥ 95	—	16.8	1.8
Trail Making Test B Seconds	90–91	—	6.5	0.0
	92–94	—	7.2	2.4
	≥ 95	—	16.8	1.8
Trail Making Test C Seconds	90–91	—	6.5	0.0
	92–94	—	6.0	2.4
	≥ 95	—	15.9	0.9
CVLT-II SF Number of words	90–91	0.8	—	—
	92–94	0.0	—	—
	≥ 95	0.9	—	—
Digit Span Test Total score	90–91	2.4	—	—
	92–94	1.2	—	—
	≥ 95	3.5	—	—

Note. BNT-Short = Boston Naming Test, Short Form (15 items); CVLT-II SF = California Verbal Learning Test-II Short Form.
[a]Percent of participants out of the total 403 in the study sample.

the three age groups, which agrees with reports of no age change or a decline with age in nonagenarians and centenarians (Kheirbek et al., 2017; Selim et al., 2005). Therefore, decline in test performance with age cannot be ascribed to differential impact of health problems in our three age groups.

The prevalence of psychoactive medication use in our group was similar to that reported in other studies of the oldest-old (Blumstein, Benyamini, Chetrit, Mizrahi, & Lerner-Geva, 2012; Wastesson, Parker, Fastbom, Thorslund, & Johnell, 2012). We observed no age difference in intake which is consistent with other reports (Wastesson et al., 2012). Therefore, we cannot ascribe the decline in test performance with age to the differential impact of psychoactive medication.

The decline in test scores with age may be related to neurodegeneration, as discussed above, but also to sensory or motor impairments. Indeed, in our sample, test non-completion due to visual or motor impairments increased with age. Cross-sectional and longitudinal studies report increased prevalence and risk of cognitive impairment in individuals with sensory impairments (Maharani et al., 2018; Mitoku, Masaki, Ogata, & Okamoto, 2016).

Scores in this study are generally comparable with other reports on cognitively normal oldest-old (Boeve et al., 2003; Fine, Kramer, Lui, Yaffe, & Study of Osteoporotic Fractures [SOF] Research Group, 2012; Iacono et al., 2014; Ivnik, Malec, Smith, Tangalos, & Petersen, 1996; Miller et al., 2015; National Alzheimer's Coordinating Center, 2017; Tombaugh, Kozak, & Rees, 1999; Weintraub et al., 2018; Zubenko, Zubenko, Maher, & Wolf, 2007). As expected, our scores are consistently higher than in studies of non-demented oldest-old that included both normal individuals and those with mild forms of cognitive impairment (Brayne, Gill, Paykel, Huppert, & O'Connor, 1995; Carrión-Baralt, Meléndez-Cabrero, Schnaider Beeri, Sano, & Silverman, 2009; Cherry et al., 2011; Elias et al., 2011; Iacono et al., 2014; Pioggiosi, Berardi, Ferrari, Quartesan, & De Ronchi, 2006; Steen, Sonn, Hanson, & Steen, 2001; Wahlin et al., 1993; Whittle et al., 2007). This is most likely due to the inclusion of individuals with mild forms of cognitive impairment in other studies as well as possible age and education differences between cohorts. Reports on centenarians and near centenarians provide lower test scores compared with our group, which could be due to higher age and the possible inclusion of cognitively impaired individuals in other cohorts (Beker et al., 2018; Davey et al., 2013, 2010; Ganz et al., 2018; Hagberg, Bauer Alfredson, Poon, & Homma, 2001; Jopp, Park, Lehrfeld, & Paggi, 2016; Miller et al., 2010).

Compared with the oldest-old population in the United States (He & Muenchrath, 2011), our sample differs little by sex (70% *vs*. 74% female), has a higher proportion of Caucasians (98.5% *vs*. 88%) and is much more highly educated (78% *vs*. 28% having more than a high school education). Although our group is not representative of other races, Caucasians are currently the overwhelming majority of the oldest-old in the United States, which makes our work relevant for most U.S. oldest-old at the present time. Our greater proportion of Caucasians is likely related to the ethnic composition of the recruitment area and highlights challenges associated with recruitment of underrepresented racial groups (Zhou et al., 2017). Our sample does not adequately represent cultural parameters, approximated by race, that are critical for test performance (Harris & Llorente, 2005). Therefore, applicability of present norms to other racial and ethnic groups is limited. In the absence of appropriate norms, it is advisable to use norms from samples most closely matching characteristics of a test-taker and to be aware of the sources of variation of test performance in different cultural groups (Ardila, 2007).

We report norms by sex and education for cognitively unimpaired oldest-old. Although in older adults quality of education (measured by reading level) (Manly, Jacobs, Touradji, Small, & Stern, 2002) or IQ score (Steinberg, Bieliauskas, Smith, Langellotti, & Ivnik, 2005) is more closely associated with neuropsychological test performance

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

Table 8. Raw MMSE and 3MS scores for participants who completed and did not complete specific neuropsychological tests

	MMSE				3MS			
Test missing	Mean for completed	Mean for missing	*t*-test	*p*-value	Mean for completed	Mean for missing	*t*-test	*p*-value
BNT-Short Total correct	28.0	27.4	−2.98	<.01	94.5	92.7	−2.85	<.01
Animal Fluency Total score	27.9	27.7	1.73	.08	94.1	94.0	−0.03	.97
Letter F Fluency Total score	27.9	29.7	−0.98	.33	94.2	93.5	−1.28	.20
CVLT-II SF Number of words	28.0	27.4	−2.23	.03	94.5	91.4	−3.67	<.01
Trail Making Test A Seconds	28.0	27.6	−1.81	.07	94.5	92.4	−3.42	<.01
Trail Making Test B Seconds	28.1	27.4	−3.77	<.01	95.0	92.3	−5.48	<.01
Trail Making Test C Seconds	28.0	27.5	−2.42	.02	94.7	92.2	−3.92	<.01
Digit Span Total score	28.1	27.6	−2.80	<.01	94.8	93.0	−3.56	<.01
Clock Drawing Total score	27.9	27.8	−0.22	.82	94.2	92.9	−1.70	.09
CERAD Construction Total score	28.0	27.5	−2.27	.02	94.5	92.3	−3.85	<.01

Note. BNT-Short = Boston Naming Test, Short Form (15 items); CVLT-II SF = California Verbal Learning Test-II Short Form; CERAD = The Consortium to Establish a Registry for Alzheimer's Disease.

than level of education, we believe that by stratifying norms by education level we likely accounted for some environmental and individual characteristics related to quality of education and IQ.

Like the majority of previously reported neuropsychological test norms, the present norms were derived from a group of participants whose cognitive status was determined cross-sectionally at the baseline evaluation. Despite our best attempt to exclude individuals with cognitive difficulties by applying clinical diagnostic criteria, a weakness of the cross-sectional approach is that individuals who go on to develop dementia may still be included into the normative sample (Sliwinski et al., 1996). In contrast, deriving norms from individuals who are cognitively normal at baseline and remain normal for several years minimizes the inclusion of individuals with preclinical dementia. This longitudinal approach to cognitive status determination likely provides greater sensitivity for the detection of cognitive impairment (Masur, Sliwinski, Lipton, Blau, & Crystal, 1994; Sliwinski et al., 1996). While attractive, this approach has several drawbacks, including the limited life expectancy in the oldest-old. However, given the potential advantages of longitudinally determined norms, we plan to explore their utility for the oldest-old.

Strengths and Limitations

This study has several notable strengths. First, we report data on one of the largest well-characterized groups of cognitively normal 90+ year olds. The large sample size made it possible to provide norms by sex and education in each of the three relatively narrow age groups. In most cases, our cell size is 50 or more participants, a desirable number for stable estimate of population mean (D'Elia, Satz, & Schretlen, 1989). Most, but not all (Ivnik et al., 1996), normative reports collapse individuals aged 90 and older into one age group or have much smaller cell sizes. With no upper age limit, we have a wider age range than age-restricted studies (Boeve et al., 2003). Second, this study, like some (Davey et al., 2010; Iacono et al., 2014; Ivnik et al., 1996; Pioggiosi et al., 2006; Tombaugh et al., 1999; Wahlin et al., 1993; Weintraub et al., 2018), but not all (Au et al., 2004; Elias et al., 2011; Fine et al., 2012) normative publications, is based on data from a study specifically designed as a cognitive aging study and uses tests well suited for the oldest-old. The tests are relatively short and involve modifications of procedures and stimuli to accommodate the sensory deficits and reduced stamina that often confound cognitive testing in old age. Third, norms are reported for tests indexing a wide range of cognition and are most frequently used by neuropsychologists. Fourth, we provide more detailed normative information, including several percentile ranges, than the majority of publications on the topic. Fifth, the detailed description of our testing procedures and scoring system facilitates data replication and tests usage. Sixth, every effort was made to collect as much testing data as possible by testing participants in their homes including traveling to other states. Seventh, cognitive status determination was based on clinical diagnostic criteria applied by trained clinicians (and not on self-report or a screening measure cutoff score) ensuring that only individuals with normal cognition were included.

We acknowledge several limitations. First, our sample represents mostly well-educated Caucasians, which limits the applicability of reported norms. Second, not all participants completed the entire test battery. Had those tests been completed, they might have affected the reported normative values. Supporting this, our analysis showed lower scores in the cognitive screening tests in individuals who did not complete individual tests compared with those who did. One of the reasons for test non-completion might be that some of the tests were more challenging than others. While we chose tests of various levels of difficulty to assess a wide range of cognitive abilities, other projects may benefit from a limited

https://doi.org/10.1017/S1355617719000122 Published online by Cambridge University Press

battery to decrease frustration, provide more valid results, and increase completion rates. Third, fixed, compared to counterbalanced, test order did not allow us to account for potential effects of the order of test administration. For instance, anxiety at the beginning and fatigue at the end of the testing may impact test performance, as may order effects such that tests administered earlier might facilitate or halt performance on subsequent tests (Franzen, Smith, Paul, & MacInnes, 1993; Llorente, Sines, Rozelle, Turcich, & Casatta, 2000). Despite the disadvantages, in *The 90+ Study* we elected to use a fixed order to ensure high completion rates of at least a few tests, given that fatigue is a major reason for test non-completion in the oldest-old. Fourth, although we strived to make our test battery comprehensive, we did not index all possible domains (e.g., fine motor skills or visual memory) to keep the battery short. Fifth, we report norms on the MMSE, 3MS, and Animal Fluency, even though these tests were used as criteria for normal cognition. We report these norms because the tests are frequently used in aging and dementia settings and their norms for the oldest-old are much needed, but the users need to be aware of the potential circularity. Sixth, the number of centenarians is limited in our group, therefore, we combined them with those aged 95 and older. We hope to provide norms for centenarians in the future as more *90+ Study* participants survive to this age.

CONCLUSIONS

Cross-sectional test norms derived from a group of cognitively normal individuals aged 90+ are instrumental in differentiating cognitively normal from impaired oldest-old. To our knowledge, this is one of the few reports on cognitive test norms derived from a large and well-characterized group of oldest-old individuals without cognitive impairment.

ACKNOWLEDGMENTS

The authors do not have any conflicts of interest to disclose. This work was supported by the National Institute on Aging (C.H.K., M.M.C., Z.A.M., A.P-H., C.W. grant R01AG021055; MBD, grant P50 AG016573). The authors thank the participants and their relatives, testers, examiners and staff of *The 90+ Study*.

SUPPLEMENTARY MATERIALS

To view supplementary material for this article, please visit https://doi.org/10.1017/S1355617719000122

REFERENCES

American Psychiatric Association. (1994). *Diagnostic and statistical manual of mental disorders*. Washington, DC: American Psychiatric Association.

Ardila, A. (2007). The impact of culture on neuropsychological test performance. In B. P. Uzzell, M. O. Ponton, & A. Ardila (Eds.), *International handbook of cross-cultural neuropsychology* (pp. 23–45). Mahwah, NJ: Lawrence Erlbaum Associates.

Au, R., Seshadri, S., Wolf, P. A., Elias, M., Elias, P., Sullivan, L., . . . D'Agostino, R. B. (2004). New norms for a new generation: Cognitive performance in the framingham offspring cohort. *Experimental Aging Research*, *30*(4), 333–358. doi:10.1080/03610730490484380

Beker, N., Sikkes, S. A. M., Hulsman, M., Schmand, B., Scheltens, P., & Holstage, H. (2018). Neuropsychological test performance of cognitively healthy centenarians: Normative data from the Dutch 100-plus study. *Journal of the American Geratrics Society*. doi:10.1111/jgs.

Blumstein, T., Benyamini, Y., Chetrit, A., Mizrahi, E. H., & Lerner-Geva, L. (2012). Prevalence and correlates of psychotropic medication use among older adults in Israel: Cross-sectional and longitudinal findings from two cohorts a decade apart. *Aging and Mental Health*, *16*(5), 636–647. doi:10.1080/13607863.2011.644262

Boeve, B., McCormick, J., Smith, G., Ferman, T., Rummans, T., Carpenter, T., . . . Petersen, R. (2003). Mild cognitive impairment in the oldest old. *Neurology*, *60*(3), 477–480.

Brayne, C., Gill, C., Paykel, E. S., Huppert, F., & O'Connor, D. W. (1995). Cognitive decline in an elderly population–A two wave study of change. *Psychological Medicine*, *25*(4), 673–683.

Carrión-Baralt, J. R., Meléndez-Cabrero, J., Schnaider Beeri, M., Sano, M., & Silverman, J. M. (2009). The neuropsychological performance of nondemented Puerto Rican nonagenarians. *Dementia and Geriatric Cognitive Disorders*, *27*(4), 353–360. doi:10.1159/000209213

Cherry, K. E., Brown, J. S., Marks, L. D., Galea, S., Volaufova, J., Lefante, C., . . . Jazwinski, S. M. (2011). Longitudinal assessment of cognitive and psychosocial functioning after hurricanes Katrina and Rita: Exploring disaster impact on middle-aged, older, and oldest-old adults. *Journal of Applied Biobehavioral Research*, *16* (3-4), 187–211. doi:10.1111/j.1751-9861.2011.00073.x

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York, NY: *Routledge Academic*.

Corrada, M. M., Brookmeyer, R., Paganini-Hill, A., Berlau, D., & Kawas, C. H. (2010). Dementia incidence continues to increase with age in the oldest old: The 90+ study. *Annals of Neurology*, *67*(1), 114–121. doi:10.1002/ana.21915

D'Elia, L., Satz, P., & Schretlen, D. (1989). Wechsler Memory Scale: A critical appraisal of the normative studies. *Journal of Clinical and Experimental Neuropsychology*, *11*(4), 551–568. doi:10.1080/01688638908400913

Davey, A., Dai, T., Woodard, J. L., Miller, L. S., Gondo, Y., Johnson, M. A., . . . Centenarian, G. (2013). Profiles of cognitive functioning in a population-based sample of centenarians using factor mixture analysis. *Experimental Aging Research*, *39*(2), 125–144. doi:10.1080/0361073X.2013.761869

Davey, A., Elias, M. F., Siegler, I. C., Lele, U., Martin, P., Johnson, M. A., . . . Poon, L. W. (2010). Cognitive function, physical performance, health, and disease: Norms from the georgia centenarian study. *Experimental Aging Research*, *36*(4), 394–425. doi:10.1080/0361073X.2010.509010

Davis, H. P., Klebe, K. J., Guinther, P. M., Schroder, K. B., Cornwell, R. E., & James, L. E. (2013). Subjective organization, verbal learning, and forgetting across the life span: From 5 to 89. *Experimental Aging Research*, *39*(1), 1–26. doi:10.1080/0361073X.2013.741956

de Azeredo Passos, V. M., Giatti, L., Bensenor, I., Tiemeier, H., Ikram, M. A., de Figueiredo, R. C., . . . Barreto, S. M. (2015).

Education plays a greater role than age in cognitive test performance among participants of the Brazilian Longitudinal Study of Adult Health (ELSA-Brasil). *BMC Neurology*, *15*, 191. doi:10.1186/s12883-015-0454-6

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan Executive Function System (D-KEFS)*. San Antonio, TX: Pearson.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (2000). *CVLT-II: California Verbal Learning Test.* (2nd ed.). San Antonio, TX: The Psychological Corporation.

Dore, G. A., Elias, M. F., Robbins, M. A., Elias, P. K., & Brennan, S. L. (2007). Cognitive performance and age: Norms from the Maine-Syracuse Study. *Experimental Aging Research*, *33*(3), 205–271. doi:10.1080/03610730701319087

Dumas, J. A., & Hartman, M. (2008). Adult age differences in the access and deletion functions of inhibition. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *15*(3), 330–357. doi:10.1080/13825580701534601

Eckert, M. A., Keren, N. I., Roberts, D. R., Calhoun, V. D., & Harris, K. C. (2010). Age-related changes in processing speed: Unique contributions of cerebellar and prefrontal cortex. *Frontiers in Human Neuroscience*, *4*, 10. doi:10.3389/neuro.09.010.2010

Elias, M. F., Dore, G. A., Goodell, A. L., Davey, A., Zilioli, M. K., Brennan, S., & Robbins, M. A. (2011). Normative data for elderly adults: The Maine-Syracuse study. *Experimental Aging Research*, *37*(2), 142–178. doi:10.1080/0361073X.2011.554511

Elias, M. F., Elias, P. K., D'Agostino, R. B., Silbershatz, H., & Wolf, P. A. (1997). Role of age, education, and gender on cognitive performance in the Framingham Heart Study: Community-based norms. *Experimental Aging Research*, *23*(3), 201–235. doi:10.1080/03610739708254281

Fastenau, P. S., Denburg, N. L., & Mauer, B. A. (1998). Parallel short forms for the Boston Naming Test: Psychometric properties and norms for older adults. *Journal of Clinical and Experimental Neuropsychology*, *20*(6), 828–834.

Fine, E. M., Kramer, J. H., Lui, L. Y., Yaffe, K., & Study of Osteoporotic Fractures (SOF) Research Group. (2012). Normative data in women aged 85 and older: Verbal fluency, digit span, and the CVLT-II short form. *The Clinical Neuropsychologist*, *26*(1), 18–30. doi:10.1080/13854046.2011.639310

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.

Franzen, M. D., Smith, S. S., Paul, D. S., & MacInnes, W. D. (1993). Order effects in the administration of the Booklet Category Test and Wisconsin Card Sorting Test. *Archives of Clinical Neuropsychology*, *8*(2), 105–110.

Gale, S. D., Baxter, L., Connor, D. J., Herring, A., & Comer, J. (2007). Sex differences on the Rey Auditory Verbal Learning Test and the Brief Visuospatial Memory Test-Revised in the elderly: Normative data in 172 participants. *Journal of Clinical and Experimental Neuropsychology*, *29*(5), 561–567. doi:10.1080/13803390600864760

Ganguli, M., Snitz, B. E., Lee, C. W., Vanderbilt, J., Saxton, J. A., & Chang, C. C. (2010). Age and education effects and norms on a cognitive test battery from a population-based cohort: The Monongahela-Youghiogheny Healthy Aging Team. *Aging and Mental Health*, *14*(1), 100–107. doi:10.1080/13607860903071014

Ganz, A. B., Beker, N., Hulsman, M., Sikkes, S., Bank, Netherlands Brain, Scheltens, P., . . . Holstege, H. (2018). Neuropathology and cognitive performance in self-reported cognitively healthy centenarians. *Acta Neuropathologica Communications*, *6*(1), 64. doi:10.1186/s40478-018-0558-5

Gasquoine, P. G. (2009). Race-norming of neuropsychological tests. *Neuropsychology Review*, *19*(2), 250–262. doi:10.1007/s11065-009-9090-5

Gladsjo, J. A., Schuman, C. C., Miller, S. W., & Heaton, R. K. (1999). *Norms for letter and category fluency: Demographic correction for age, education, and ethnicity*. Odessa, FL: PAR.

Graham, J. E., Rockwood, K., Beattie, B. L., Eastwood, R., Gauthier, S., Tuokko, H., & McDowell, I. (1997). Prevalence and severity of cognitive impairment with and without dementia in an elderly population. *Lancet*, *349*(9068), 1793–1796. doi:10.1016/S0140-6736(97)01007-6

Hafkemeijer, A., Altmann-Schneider, I., de Craen, A. J., Slagboom, P. E., van der Grond, J., & Rombouts, S. A. (2014). Associations between age and gray matter volume in anatomical brain networks in middle-aged to older adults. *Aging Cell*, *13*(6), 1068–1074. doi:10.1111/acel.12271

Hagberg, B., Bauer Alfredson, B., Poon, L. W., & Homma, A. (2001). Cognitive functioning in centenarians: A coordinated analysis of results from three countries. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *56*(3), P141–P151.

Hall, J. R., Vo, H. T., Johnson, L. A., Wiechmann, A., & O'Bryant, S. E. (2012). Boston Naming Test: Gender differences in older adults with and without Alzheimer's dementia. *Psychology*, *3*(6), 485–488.

Harada, C. N., Natelson Love, M. C., & Triebel, K. L. (2013). Normal cognitive aging. *Clinics in Geriatric Medicine*, *29*(4), 737–752. doi:10.1016/j.cger.2013.07.002

Harris, J. G., & Llorente, A. M. (2005). Cultural consideration in the use of the Wechsler Intelligence Scale for Children - Fourth Edition (WISC-IV). In A. Prifitera, D. H. Saklofske, & L. G. Weiss (Eds.), *WISC-IV clinical use and interpretation: Scientist-practitioner perspectives* (pp. 382–413). Burlington, VT: Elsevier Academic.

He, W., & Muenchrath, M. (2011). *ACS-17 90+ in the United States: 2006-2008. American Community Survey Reports*. Retrieved from https://www2.census.gov/library/publications/2011/acs/acs-17.pdf

Heaton, R. K., Miller, S. W., Taylor, M. J., & Grant, I. (2004). *Revised comprehensive norms for an expanded Halstead-Reitan battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults*. Lutz, FL: PAR.

Hogervorst, E., Rahardjo, T. B., Jolles, J., Brayne, C., & Henderson, V. W. (2012). Gender differences in verbal learning in older participants. *Journal of Aging and Health*, *8*(5), 1–15.

Iacono, D., Resnick, S. M., O'Brien, R., Zonderman, A. B., An, Y., Pletnikova, O., . . . Troncoso, J. C. (2014). Mild cognitive impairment and asymptomatic Alzheimer disease subjects: Equivalent β-amyloid and tau loads with divergent cognitive outcomes. *Journal of Neuropathology and Experimental Neurology*, *73*(4), 295–304. doi:10.1097/NEN.0000000000000052

Ivnik, R. J., Malec, J. F., Smith, G. E., Tangalos, E. G., & Petersen, R. C. (1996). Neuropsychological Tests' Norms Above Age 55: COWAT, BNT, MAE Token, WRAT-R Reading, AMNART, STROOP, TMT, and JLO. *The Clinical Neuropsychologist*, *10*(3), 262–278.

Jopp, D. S., Park, M. K., Lehrfeld, J., & Paggi, M. E. (2016). Physical, cognitive, social and mental health in near-centenarians and centenarians living in New York City: Findings from the

Fordham Centenarian Study. *BMC Geriatrics*, *16*, 1. doi:10.1186/s12877-015-0167-0

Kheirbek, R. E., Fokar, A., Shara, N., Bell-Wilson, L. K., Moore, H. J., Olsen, E., . . . Llorente, M. D. (2017). Characteristics and incidence of chronic illness in community-dwelling predominantly male U.S. veteran centenarians. *Journal of the American Geriatrics Society*, *65*(9), 2100–2106. doi:10.1111/jgs.14900

Koenig, A. M., Bhalla, R. K., & Butters, M. A. (2014). Cognitive functioning and late-life depression. *Journal of the International Neuropsychological Society*, *20*(5), 461–467. doi:10.1017/S1355617714000198

Lee, S. J., Go, A. S., Lindquist, K., Bertenthal, D., & Covinsky, K. E. (2008). Chronic conditions and mortality among the oldest old. *American Journal of Public Health*, *98*(7), 1209–1214. doi:10.2105/AJPH.2007.130955

Legdeur, N., Binnekade, T. T., Otten, R. H., Badissi, M., Scheltens, P., Visser, P. J., & Maier, A. B. (2017). Cognitive functioning of individuals aged 90 years and older without dementia: A systematic review. *Ageing Research Reviews*, *36*, 42–49. doi:10.1016/j.arr.2017.02.006

Liu, H., Yang, Y., Xia, Y., Zhu, W., Leak, R. K., Wei, Z., . . . Hu, X. (2017). Aging of cerebral white matter. *Ageing Research Reviews*, *34*, 64–76. doi:10.1016/j.arr.2016.11.006

Llorente, A. M., Sines, M. C., Rozelle, J. C., Turcich, M. R., & Casatta, A. (2000). Effects of test administration order on children's neuropsychological performance: Emerging one-word expressive and receptive language skills. *The Clinical Neuropsychologist*, *14*(2), 162–172. doi:10.1076/1385-4046(200005)14:2;1-Z;FT162

Maharani, A., Dawes, P., Nazroo, J., Tampubolon, G., Pendleton, N., & Sense-Cog WP1 Group. (2018). Visual and hearing impairments are associated with cognitive decline in older people. *Age and Ageing*. doi:10.1093/ageing/afy061

Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society*, *8*(3), 341–348.

Masur, D. M., Sliwinski, M., Lipton, R. B., Blau, A. D., & Crystal, H. A. (1994). Neuropsychological prediction of dementia and the absence of dementia in healthy elderly persons. *Neurology*, *44*(8), 1427–1432.

Melikyan, Z. A., Greenia, D. E., Corrada, M. M., Hester, M. M., Kawas, C. H., & Grill, J. D. (2018). Recruiting the oldest-old for clinical research. *Alzheimer Disease and Associated Disorders*. doi:10.1097/WAD.0000000000000260

Miller, I. N., Himali, J. J., Beiser, A. S., Murabito, J. M., Seshadri, S., Wolf, P. A., & Au, R. (2015). Normative data for the cognitively intact oldest-old: The Framingham Heart Study. *Experimental Aging Research*, *41*(4), 386–409. doi:10.1080/0361073X.2015.1053755

Miller, L. S., Mitchell, M. B., Woodard, J. L., Davey, A., Martin, P., Poon, L. W., . . . Siegler, I. C. (2010). Cognitive performance in centenarians and the oldest old: Norms from the Georgia Centenarian Study. *Neuropsychology, Development, and Cognition. Section B, Aging, Neuropsychology and Cognition*, *17*(5), 575–590. doi:10.1080/13825585.2010.481355

Mitoku, K., Masaki, N., Ogata, Y., & Okamoto, K. (2016). Vision and hearing impairments, cognitive impairment and mortality among long-term care recipients: A population-based cohort study. *BMC Geriatrics*, *16*, 112. doi:10.1186/s12877-016-0286-2

Morimoto, S. S., & Alexopoulos, G. S. (2013). Cognitive deficits in geriatric depression: Clinical correlates and implications for current and future treatment. *The Psychiatric Clinics of North America*, *36*(4), 517–531. doi:10.1016/j.psc.2013.08.002

Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, *43*(11), 2412–2414.

Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., . . . Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, *39*(9), 1159–1165.

National Alzheimer's Coordinating Center. (2017). Means and standard deviations for the UDS3 neuropsychological battery in cognitively normal participants - March 2017 - NACC. Retreived from https://www.alz.washington.edu/WEB/UDS3means.pdf https://www.alz.washington.edu/WEB/UDS3means.pdf

Nosraty, L., Sarkeala, T., Hervonen, A., & Jylhä, M. (2012). Is there successful aging for nonagenarians? The vitality 90+ study. *Journal of Aging Research*, *2012*, 868797. doi:10.1155/2012/868797

Paganini-Hill, A., Ross, R. K., & Henderson, B. E. (1986). Prevalence of chronic disease and health practices in a retirement community. *Journal of Chronic Disease*, *39*(9), 699–707.

Pfeffer, R. I., Kurosaki, T. T., Harrah, C. H., Chance, J. M., & Filos, S. (1982). Measurement of functional activities in older adults in the community. *Journal of Gerontology*, *37*(3), 323–329.

Pioggiosi, P. P., Berardi, D., Ferrari, B., Quartesan, R., & De Ronchi, D. (2006). Occurrence of cognitive impairment after age 90: MCI and other broadly used concepts. *Brain Research Bulletin*, *68*(4), 227–232. doi:10.1016/j.brainresbull.2005.06.039

Rabin, L. A., Paolillo, E., & Barr, W. B. (2016). Stability in test-usage practices of clinical neuropsychologists in the United States and Canada over a 10-year period: A follow-up survey of INS and NAN members. *Archives of Clinical Neuropsychology*, *31*(3), 206–230. doi:10.1093/arclin/acw007

Rasmusson, X., Zonderman, A., Kawas, C., & Resnick, S. M. (1998). Effects of age and dementia on Trail Making Test. *The Clinical Neuropsychologist*, *12*(2), 169–178.

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

Rosselli, M., Tappen, R., Williams, C., & Salvatierra, J. (2006). The relation of education and gender on the attention items of the Mini-Mental State Examination in Spanish speaking Hispanic elders. *Archives of Clinical Neuropsychology*, *21*(7), 677–686. doi:10.1016/j.acn.2006.08.001

Saykin, A. J., Gur, R. C., Gur, R. E., Shtasel, D. L., Flannery, K. A., Mozley, L. H., . . . Mozley, P. D. (1995). Normative neuropsychological test performance: Effects of age, education, gender and ethnicity. *Applied Neuropsychology*, *2*(2), 79–88. doi:10.1207/s15324826an0202_5

Selim, A. J., Fincke, G., Berlowitz, D. R., Miller, D. R., Qian, S. X., Lee, A., . . . Kazis, L. E. (2005). Comprehensive health status assessment of centenarians: Results from the 1999 large health survey of veteran enrollees. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *60*(4), 515–519.

Sliwinski, M., Lipton, R. B., Buschke, H., & Stewart, W. (1996). The effects of preclinical dementia on estimates of normal cognitive functioning in aging. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, *51*(4), P217–P225.

Steen, G., Sonn, U., Hanson, A. B., & Steen, B. (2001). Cognitive function and functional ability. A cross-sectional and longitudinal study at ages 85 and 95 in a non-demented population. *Aging (Milano)*, *13*(2), 68–77.

Steinberg, B. A., Bieliauskas, L. A., Smith, G. E., Langellotti, C., & Ivnik, R. J. (2005). Mayo's older americans normative studies: Age- and IQ-adjusted norms for the Boston Naming Test, the MAE Token Test, and the Judgment of Line Orientation Test. *The Clinical Neuropsychologist*, *19*(3-4), 280–328. doi:10.1080/13854040590945229

Stern, Y. (2012). Cognitive reserve in ageing and Alzheimer's disease. *Lancet Neurology*, *11*(11), 1006–1012. doi:10.1016/S1474-4422(12)70191-6

Teng, E. L., & Chui, H. C. (1987). The Modified Mini-Mental State (3MS) examination. *The Journal of Clinical Psychiatry*, *48*(8), 314–318.

Tombaugh, T. N., Kozak, J., & Rees, L. (1999). Normative data stratified by age and education for two measures of verbal fluency: FAS and animal naming. *Archives of Clinical Neuropsychology*, *14*(2), 167–177.

United Nations Department of Economic and Social Affairs Population Division. (2017). World Population Prospects: The 2017. Revision. Retrieved from https://www.un.org/development/desa/publications/world-population-prospects-the-2017-revision.html

Wahlin, A., Bäckman, L., Mäntylä, T., Herlitz, A., Viitanen, M., & Winblad, B. (1993). Prior knowledge and face recognition in a community-based sample of healthy, very old adults. *Journal of Gerontology*, *48*(2), P54–P61.

Wastesson, J. W., Parker, M. G., Fastbom, J., Thorslund, M., & Johnell, K. (2012). Drug use in centenarians compared with nonagenarians and octogenarians in Sweden: A nationwide register-based study. *Age and Ageing*, *41*(2), 218–224. doi:10.1093/ageing/afr144

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale - Third Edition (WAIS-III)*. San Antonio, TX: Psychological Corporation.

Weintraub, S., Besser, L., Dodge, H. H., Teylan, M., Ferris, S., Goldstein, F. C., . . . Morris, J. C. (2018). Version 3 of the Alzheimer Disease Centers' Neuropsychological Test Battery in the Uniform Data Set (UDS). *Alzheimer Disease and Associated Disorders*, *32*(1), 10–17. doi:10.1097/WAD.0000000000000223

Weisenbach, S. L., Boore, L. A., & Kales, H. C. (2012). Depression and cognitive impairment in older adults. *Current Psychiatry Reports*, *14*(4), 280–288. doi:10.1007/s11920-012-0278-7

Whittle, C., Corrada, M. M., Dick, M., Ziegler, R., Kahle-Wrobleski, K., Paganini-Hill, A., & Kawas, C. (2007). Neuropsychological data in nondemented oldest old: The 90+ Study. *Journal of Clinical and Experimental Neuropsychology*, *29*(3), 290–299. doi:10.1080/13803390600678038

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, V. O. (1982-1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *17*(1), 37–49.

Zhou, Y., Elashoff, D., Kremen, S., Teng, E., Karlawish, J., & Grill, J. D. (2017). African Americans are less likely to enroll in preclinical Alzheimer's disease clinical trials. *Alzheimers Dement (N Y)*, *3*(1), 57–64. doi:10.1016/j.trci.2016.09.004

Zubenko, G. S., Zubenko, W. N., Maher, B. S., & Wolf, N. S. (2007). Reduced age-related cataracts among elderly persons who reach age 90 with preserved cognition: A biomarker of successful aging? *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *62*(5), 500–506.