ORIGINAL ARTICLE

Version 3 of the Alzheimer Disease Centers' Neuropsychological Test Battery in the Uniform Data Set (UDS)

Sandra Weintraub, PhD,*† Lilah Besser, PhD,‡ Hiroko H. Dodge, PhD,§‖
Merilee Teylan, MPH,‡ Steven Ferris, PhD✠,¶ Felicia C. Goldstein, PhD,#
Bruno Giordani, PhD,‖ Joel Kramer, PsyD,** David Loewenstein, PhD,††
Dan Marson, PhD,‡‡ Dan Mungas, PhD,§§ David Salmon, PhD,‖‖
Kathleen Welsh-Bohmer, PhD,¶¶ Xiao-Hua Zhou, PhD,‡
Steven D. Shirk, PhD,## Alireza Atri, MD, PhD,***†††‡‡‡
Walter A. Kukull, PhD, MS,‡ Creighton Phelps, PhD,§§§
and John C. Morris, MD‖‖‖

Introduction: The neuropsychological battery of the Uniform Data Set (UDSNB) was implemented in 2005 by the National Institute on Aging (NIA) Alzheimer Disease Centers program to measure cognitive performance in dementia and mild cognitive impairment due to Alzheimer Disease. This paper describes a revision, the UDSNB 3.0.

Methods: The Neuropsychology Work Group of the NIA Clinical Task Force recommended revisions through a process of due diligence to address shortcomings of the original battery. The UDSNB 3.0 covers episodic memory, processing speed, executive function, language, and constructional ability. Data from 3602 cognitively normal participants in the National Alzheimer Coordinating Center database were analyzed.

Results: Descriptive statistics are presented. Multivariable linear regression analyses demonstrated score differences by age, sex, and education and were also used to create a normative calculator available online.

Discussion: The UDSNB 3.0 neuropsychological battery provides a valuable non proprietary resource for conducting research on cognitive aging and dementia.

Key Words: dementia, neuropsychological test, cognition, UDS

(*Alzheimer Dis Assoc Disord* 2018;32:10–17)

Since 2005, the University of Washington's National Alzheimer's Coordinating Center (NACC) has collected the Uniform Data Set (UDS) on participants from over 30 past and present US Alzheimer's Disease Centers (ADC).

The UDS consists of data collection protocols used systematically on participants enrolled into the Clinical Cores of each ADC.[1,2] Participants with clinical diagnoses of normal cognition (NC), mild cognitive impairment (MCI) and dementia of various etiologies including Alzheimer disease (AD) are recruited, enrolled, and followed annually. Consent is obtained at the individual ADCs, as approved by individual Institutional Review Boards (IRBs), and the University of Washington's IRB has approved the sharing of deidentified UDS data. The UDS data, include demographics, medical history, medication use, clinical and neurological examination findings, measures of function and behavior, clinical ratings of dementia severity [eg, clinical dementia rating (CDR)[3]], and neuropsychological test scores. Systematic guidelines for clinical diagnosis are based on the most up to date published diagnostic research criteria.[1,4]

All UDS data collection instruments were constructed with the guidance and approval of the Clinical Task Force (CTF), a group originally constituted by the National Institute on Aging (NIA) to develop standardized methods for collecting longitudinal data that would encourage and support collaboration across the ADCs.[1,5] As of December 1, 2016, the NACC database contained data on 34,748 UDS participants from past and present ADCs.

ADCs used the first version of the UDSNB starting in September 2005, and in February 2008, a second version, UDSNB 2.0, was implemented with slight revisions to instructions and data collection forms. Tests in the original version of the battery, used until March 2015, were chosen to capture the continuum of cognitive decline from NC through AD dementia, incorporating relevant domains described in detail previously.[2] An online calculator was developed to aid in scoring.[6]

In 2010, encouraged by recognition of the growing importance of diagnostic biomarkers[7] and the identification of preclinical stages of AD,[8] the Neuropsychology Work Group, a committee to review the UDSNB 2.0 and make recommendations for future data collection, was convened. This paper describes the rationale and procedures for the development of UDSNB 3.0 and provides normative test scores for a cohort of cognitively normal individuals from the NACC database.

METHODS

Rationale and Procedures for Battery Design and Test Selection

The CTF and Neuropsychology Work Group outlined the rationale for change. First, in longitudinal follow-up, healthy controls showed practice effects, especially for the memory task, even reciting the story before administration on visits subsequent to baseline. UDSNB 2.0 measures were published tests, increasing the potential for multiple exposures either through clinical practice or in ancillary research conducted at the ADCs. Licensing costs and restrictions on sharing these instruments with intramural, extramural, and international researchers also created challenges for collaboration. Furthermore, the importance of early detection required instruments that would be sensitive to earlier stages of cognitive decline or even "preclinical" states. Finally, UDSNB 2.0 lacked tests of visuospatial functions and nonverbal memory, both of which can constitute areas of early decline, particularly in those with Lewy body

disease[9,10] or those with the posterior cortical atrophy variant of AD.[11–13] Therefore, we created novel tests to address some of the shortcomings of the existing battery, while at the same time having a mechanism for preserving longitudinal continuity with previous data.

The Work Group included members from several ADCs to ensure multicenter representation, and exofficio members from the NACC and the NIA. The group conducted weekly or monthly conference calls, as needed, and in-person meetings 1 to 2 times per year to outline a strategy to assess options for change. This included considering different platforms for testing (paper-and-pencil vs. computer), considering whether or not to change the types of constructs tested with the UDSNB 3.0, and evaluating existing instruments for inclusion into the new battery. Criteria were developed to aid in decision-making. Early on, after reviewing a library of potential tests, the group decided to adopt nonproprietary measures to allow the ADCs to freely share the battery with collaborators. Moreover, early on, it was decided to postpone computerized testing, as the field was rapidly evolving with increasingly sophisticated technology.

The UDSNB 2.0 faced the serious issue of a new battery disrupting the longitudinal follow-up of participants tested with the initial battery since 2008. Therefore, the decision was made after careful review by the Work Group and the CTF and presentations to the centers, to model the new battery on the old one and to drop or replace existing measures. Digit Symbol from the WAIS-R was dropped, whereas the trail making tests and category list generation tests (animals, fruits, and vegetables) were retained. In addition, 4 measures were replaced with similar measures developed previously by several of the centers and tested in published research studies. The section below describes the instruments.

Materials, data recording forms, and a manual for administration and scoring were created and revised with feedback from the centers. After a brief period of pilot data collection with the new instruments to refine the instructions and address any questions about administration and scoring we made additional revisions and conducted a larger pilot study (N = 935) that compared the UDSNB 2.0 and 3.0 versions in individuals divided into 4 groups based on their Mini Mental State Examination (MMSE)[14] scores (26 to 30, 21 to 25, 16 to 20, 10 to 15) in a "crosswalk" study.[15] The pairs of scores for the original and corresponding replacement tests were compared using equipercentile equating, and the analyses provided a crosswalk of equivalent test scores between the original and replacement tests [eg, a score of 15 on the Montreal Cognitive Assessment (MoCA) is equivalent to a score of 21 on the MMSE]. The results of the crosswalk study provided good evidence for relatively reliable equivalence across both measures and that the chosen tests were reasonable replacements for the older tests. Crosswalk scores could also assist in making longitudinal comparisons.

Selection of Tests for the UDS Neuropsychological Battery 3.0

The Work Group recommended replacing the MMSE with the MoCA,[16,17] Logical Memory[18] immediate and delayed with the Craft Story 21 immediate and delayed recall[19]; digit span forward and backward with the number span forward and backward test; and the Boston naming test (BNT) with the multilingual naming test (MINT).[20] Each decision was based on the rationale outlined below.

General Cognitive Measure

The MoCA[16,17] was selected to replace the MMSE as a measure of overall cognitive impairment. Factors influencing this decision included the fact that the MOCA is more difficult than the MMSE as demonstrated in studies showing lower MoCA than MMSE scores in the same samples[21] and, hence, more likely to detect subtle cognitive deficits. Furthermore, floor and ceiling effects are less common with the MoCA, which also allows for a broader range of scores in MCI samples than does the MMSE.[22] Therefore, the MoCA is more appropriate than the MMSE for detecting early cognitive decline. The MoCA has been validated in white[23] and African American[24] groups. A disadvantage of the MoCA is that it can yield lower scores in diverse healthy population-based samples.[25] However, an abbreviated version reportedly demonstrated predictive ability with respect to diagnosis of MCI in a low-education, illiterate sample.[26] In another study, MoCA was more sensitive to MCI and discriminated MCI from other samples better than the MMSE.[27,28] MoCA scores have also been shown to correlate with the Activities of Daily Living Questionnaire[29] a measure of functional integrity in dementia.[30] The MoCA has the further advantage of yielding not only a total score (overall measure) but also index scores based on individual items tapping domains of attention, retentive memory, orientation, language, and executive function.[31]

The MoCA requires about 10 minutes to administer and yields a total score of 30 and the above-mentioned domain index scores. The index scores (not included in the present report) offer the potential to identify early dementia profiles of clinical dementia subtypes such as behavioral variant frontotemporal dementia and primary progressive aphasia. The memory index score has been shown to be especially predictive of decline from amnestic MCI to AD dementia.[31] The paper-and-pencil version of the MoCA has been translated into multiple languages and dialects within languages[32] and is freely available (www.mocatest.org/). The NACC was given permission to use it for 25 years without royalties or restrictions on sharing the test with collaborators.

Development of Domain-specific Neuropsychological Tests

Episodic Memory Tests

Memory loss is the hallmark symptom of the most common clinical dementia syndrome associated with AD.[33] Early studies of AD dementia emphasized the importance of measures of episodic memory, such as word list learning and story recall, in the evaluation for dementia. The group had decided on a story memory test, as most ADCs were already using Logical Memory, immediate, and delayed recall conditions.

Craft and colleagues had designed multiple forms of a story recall test similar to Logical Memory in a study of the impact of insulin on cognition in mild AD dementia.[19,34] The complete set of 22 stories had previously been tested for equivalence in a diverse sample of college age adults who were administered all of the stories in counter balanced order in the laboratory of Andrew Saykin (personal oral communication) and provided to the Work Group for consideration. Additional data on alternate sets of stories were included in published studies of patients undergoing systemic chemotherapy for treatment of breast cancer as well as individuals with traumatic brain injury and healthy controls.[35–38] In a pilot study to determine the equivalence of 22 stories in middle-aged and older adults the Work Group determined that 3 stories offered the greatest relationship to Logical Memory and to one another. These 3 were reviewed by the work group and 1 was chosen for its content relevance to a diverse population, "Craft Story 21."

Scoring of Logical Memory allows several acceptable responses for each item recalled. Following the protocol from Craft et al,[34] items were scored in a similar manner to Logical Memory (paraphrase score) but another score was also calculated (verbatim score), allocating a point for each item recalled exactly as delivered in the story. The verbatim score (not included in the present report) was intended to serve as potentially more sensitive than the paraphrase score in detecting very early memory decline.

Finally, we introduced a novel measure of nonverbal memory, a function not previously included in the UDSNB 2.0. Following the copy of the Benson complex[39] figure (see the Visuospatial Test section) delayed figure reproduction was tested.

Language Tests

The 32-item MINT[20,40] was selected to replace the short BNT. The MINT was originally developed to test naming in 4 languages, English, Spanish, Hebrew, and Mandarin Chinese, taking care to equate the level of difficulty of items across languages. The BNT was developed in New England and designed for American English speakers and contains items that either have no equivalent word or different frequencies of usage in other languages. The MINT is sensitive to naming impairment in AD.[20]

Word fluency is measured with semantic and letter word list generation tests. The former were part of UDSNB 2.0, whereas 2 letter generation tasks were added ("F" and "L") for UDSNB 3.0. Each task requires 60 seconds and correct items are totaled. Note is made of errors and rule violations.

Visuospatial Tests

The UDSNB 2.0 did not contain a visuospatial test. Visuospatial symptoms emerge in later stages of amnestic dementia due to AD but also may appear early in the clinical syndromes of posterior cortical atrophy and dementia associated with cortical Lewy body disease. The Benson complex figure[39] was added as a test of constructional ability (Copy condition). Figural elements are scored for presence and placement. Reproduction is tested after a delay to measure retentive memory (see the Episodic Memory section). Comparison between patients with clinical dementia of the Alzheimer type and frontotemporal dementia showed distinctive profiles of performance and associations with frontal and parietal cortical atrophy regions in the groups.[39,41]

Immediate Attention, Working Memory, Executive Attention Tests

Immediate attention span is commonly tested with Digit Span.[18] For studies requiring multiple forms to reduce practice effects, a series of number sets was randomly generated to provide alternatives to the digit span test (Joel Kramer lab, personal oral communication). The number spans for the UDS task were randomly generated with the restriction that no digit would be adjacent to a digit that was one higher or one lower (eg, a "7" would not be succeeded or preceded by a 6 or 8). Every attempt was also made to exclude sequences that contained area codes. The number span is the longest list recalled. The total number of trials

administered up to failure on 2 trials at 1 length is also recorded. Backward span is a measure of working memory. The trail making tests were retained from the UDSNB 2.0 to measure processing speed (part A) and executive attention (part B).

Study Sample

This report is based on analyses of UDS data submitted to NACC by the ADCs between March 15, 2015 and November 30, 2016. The sample was restricted to individuals who received the UDSNB 3.0 and at that visit had a clinical diagnosis of NC and a global CDR score of 0. If a participant had received UDSNB 3.0 more than once, data were included from only the first administration. Although some participants' scores on the UDSNB 3.0 seemed to be outside the range of normal scores (eg, MoCA score of 9), we chose not to remove any participants from the descriptive analyses because normalcy was not defined by the tests. Therefore, we describe the full range of scores in those with a clinical diagnosis of NC and a global CDR = 0.

Data, Analyses, Normative Calculator

First we describe the demographics of the sample (age, education, and sex). The mean, median, 25th and 75th percentiles, and ranges of scores for the overall sample are presented. Histograms are provided for each of the tests to illustrate the distribution of scores in the overall sample. The mean scores and SDs for each test are provided by age divided into 5 groups (< 60, 60 to 69, 70 to 79, 80 to 89, ≥ 90 y) and education, divided into 4 groups (≤ 12, 13 to 15, 16, ≥ 17 y). Unadjusted linear regression analyses tested for differences by age or education group. Finally, we ran linear regression models to estimate the effect of age (continuous), sex, and education (continuous) on each neuropsychological measure. Adjusted linear regression models were first run with either age, sex, or education predicting the neuropsychological test score (data not shown), and then multivariable models were run with all 3 demographics included in the model.

We developed a calculator for the UDSNB 3.0 tests based on previously published methods used to produce the calculator for UDSNB 2.0 tests.[6] Although our descriptive analyses focused on all participants meeting our eligibility criteria, for the normative calculator, we excluded a handful of participants who performed 5 SDs outside of the mean on any particular test to improve the distribution of residuals and better satisfy model assumptions. This restriction resulted in excluding the following participants from the regression analyses: 5 participants from the analysis of the MoCA, 4 participants from the analysis of the Benson complex figure copy, 16 participants from the analysis of the trail making part A, and 5 participants from the analysis of the MINT.

RESULTS

The sample included 3602 cognitively normal participants over age 60 receiving the UDSNB 3.0 (Table 1). The majority of the sample (65%) were women and were between 70 and 89 years of age (67%) and highly educated (69%). These analyses did not divide the sample by race, as most participants in the sample were white (83%), with an additional 14% African American, and 3% other race, reflecting the overall distribution of these groups within the ADCs receiving the UDS.

Means, 25th, 50th, and 75th percentile, and score ranges for each test in the overall sample are reported in

TABLE 1. Sample Distribution by Sex, Age, and Education

Age (y)	Education	Male	Female	Total
< 60	≤ 12	13	23	36
	13-15	23	48	71
	16	34	69	103
	17+	44	76	120
	Missing	6	6	12
60-69	≤ 12	32	65	97
	13-15	47	142	189
	16	102	164	266
	17+	158	293	451
	Missing	3	7	10
70-79	≤ 12	45	130	175
	13-15	61	159	220
	16	110	201	311
	17+	274	407	681
	Missing	1	7	8
80-89	≤ 12	27	77	104
	13-15	26	103	129
	16	59	112	171
	17+	151	177	328
	Missing	1	1	2
≥ 90	≤ 12	4	27	31
	13-15	6	11	17
	16	10	15	25
	17+	19	26	45
	Missing	0	0	0
Grand total		1256	2346	3602

As of December 2016.

Table 2. Histograms demonstrate whether the distribution of test scores were approximately normal (Fig. 1). Tests with an approximately normal distribution of scores included Craft Story immediate and delayed (paraphrase and verbatim), number span forward and backward (total correct trials and longest span), the letter list generation task (F-words and L-words), and the Benson complex figure recall. Scores on the MoCA, MINT, and copy condition of the Benson complex figure copy were highly skewed due to ceiling effects. However, the MoCA seems to be less affected by ceiling effects than the MMSE.[2]

Table 3 shows the means and SDs by each measure across the 5 age groups, and Table 4 shows means and SDs for the 4 education groups (Table 5). In the multivariable regression analyses (Table 5), women performed statistically significantly ($P < 0.01$) better than men on the Craft story immediate and delayed, verbal fluency phonemic test, and vegetables list generation, but worse on the Benson copy figure recall, number span forward, and MINT (Table 5). Women and men performed similarly, without statistically significant differences, on the Benson complex figure copy, number span backward, animal list generation, and trail making parts A and B. Increasing age was associated with worse scores and increasing years of education was associated with better scores on all of the tests ($P < 0.01$).

For the data to be useful in characterizing research participants, a calculator was created to indicate the level of performance on each measure. The calculator uses the intercepts, regression coefficients, and root mean square errors resulting from the regression analyses described above to calculate unadjusted and adjusted Z-scores for individuals of a particular sex, age, and/or education level. The root mean square error is the square root of the

TABLE 2. Summary Statistics for Clinically Cognitively Normal UDS Participants

UDS Version 3 Neuropsychological Test*	Domain	Maximum Score	N	Mean (SD)	Q_{25}, Q_{50}, Q_{75}	Range
				Sample's Scores		
MoCA, total score	Dementia severity	30	3581	26.3 (2.8)	25, 27, 28	9-30
Craft Story 21 recall immed. verbatim, total units	Memory	44	3552	21.9 (6.6)	17, 22, 27	0-38
Craft Story 21 recall immed. paraphrase, total units	Memory	25	3552	16.1 (4.1)	14, 17, 19	0-25
Craft Story 21 recall delay. verbatim, total units	Memory	44	3550	19.1 (6.7)	14, 19, 24	0-40
Craft Story 21 recall delay. Paraphrase, total units	Memory	25	3550	15.1 (4.3)	12, 16, 18	0-25
Benson complex figure copy, total score	Visuospatial	17	3584	15.6 (1.4)	15, 16, 17	0-17
Benson complex figure recall, total score	Visuospatial/Memory	17	3576	11.2 (3.1)	9, 12, 13	0-17
Number span test forward, total correct trials	Attention	14	3591	8.3 (2.3)	7, 8, 10	2-14
Number span test forward, longest span	Attention	9	3591	6.7 (1.3)	6, 7, 8	3-9
Number span test backward, total correct trials	Attention	14	3590	7.2 (2.2)	6, 7, 9	0-14
Number span test backward, longest span	Attention	8	3590	5.1 (1.3)	4, 5, 6	0-8
MINT, total score	Lang. naming	32	3564	30.0 (2.4)	29, 31, 32	0-32
Phonemic test, F-words total in 60 s	Lang. verbal fluency	40	3589	15.1 (4.7)	12, 15, 18	1-35
Phonemic test, L words total in 60 s	Lang. verbal fluency	40	3574	14.2 (4.5)	11, 14, 17	0-35
Phonemic test, total F-words and L-words	Lang. verbal fluency	80	3572	29.2 (8.6)	23, 29, 35	1-64
Animals list generation, total in 60 s	Lang. category fluency	77	3596	21.4 (5.7)	17, 21, 25	0-49
Vegetables list generation, total in 60 s	Lang. category fluency	77	3590	15.1 (4.3)	12, 15, 18	1-35
Trail making test part A, time (s)	Processing speed	150	3583	30.9 (13.4)	22, 28, 36	9-150
Trail making test part B, time (s)	Executive function	300	3561	82.2 (46.3)	54, 70, 95	13-300

As of December 2016.

*Higher scores indicate better scores except for the trail making test parts A and B.

Delay indicates delayed; Immed, immediate; Lang, language; Max, maximum; MINT, multilingual naming test; MoCA, Montreal Cognitive Assessment; UDS, Uniform Data Set.

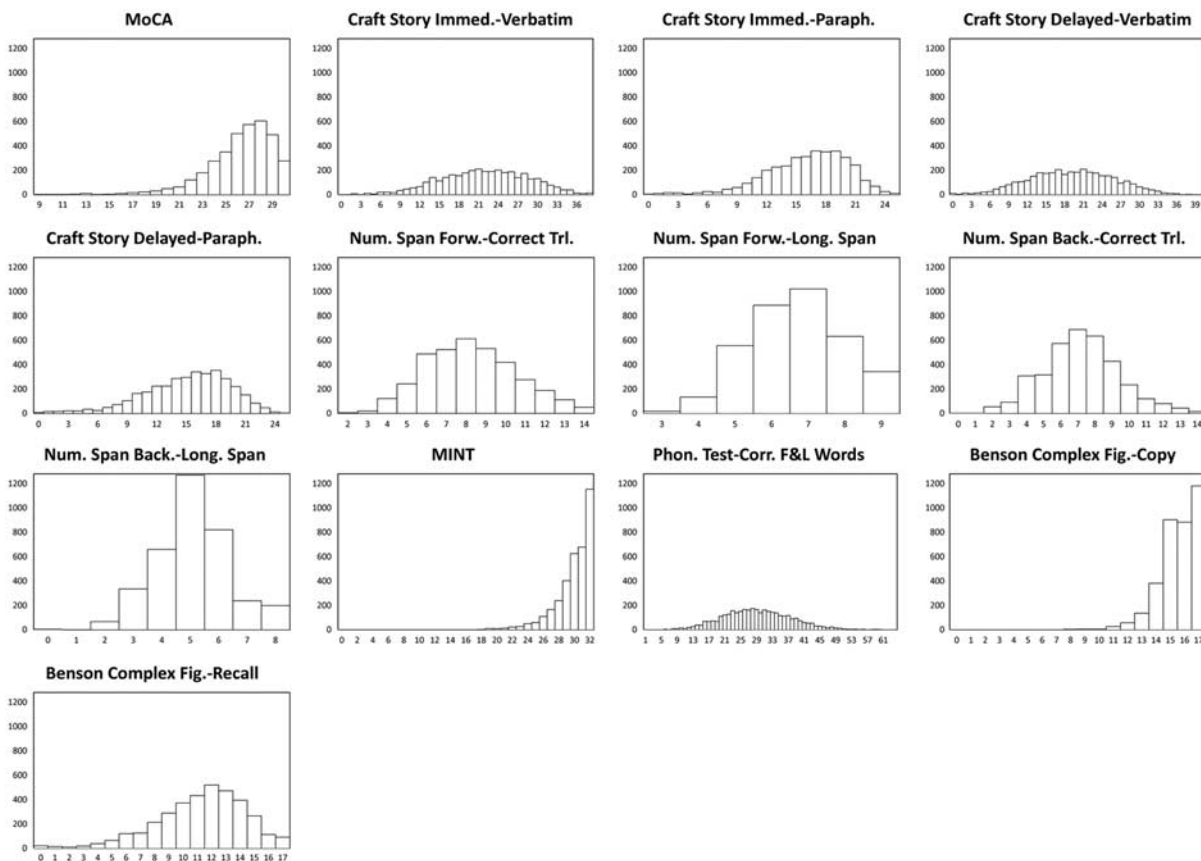**FIGURE 1.** Histograms showing score distributions for each measure on the UDSNB 3.0. From these graphs, many of the measures have a normal or near normal distribution, with the exception of the MoCA total score, the score for the copy of the Benson complex figure, and the total score for the MINT. Immed. indicates immediate; MINT, multilingual naming test; MoCA, Montreal Cognitive Assessment; Paraph., paraphrase.

TABLE 3. Mean Neuropsychological Test Scores by Age Group

UDS Version 3 Neuropsychological Test*	Mean (SD)				
	< 60 y	60-69 y	70-79 y	80-89 y	≥ 90 y
Montreal Cognitive Assessment, total score	27.5 (2.1)	26.9 (2.4)	26.3 (2.7)	25.3 (3.0)	23.8 (3.5)
Craft Story 21 recall immed. verbat., total units	23.1 (6.8)	23.2 (6.4)	22.0 (6.5)	20.0 (6.6)	17.8 (6.4)
Craft Story 21 recall immed. paraph., total units	16.9 (4.1)	16.9 (3.9)	16.2 (4.0)	15.0 (4.2)	13.3 (4.4)
Craft Story 21 recall delay. verbat. total units	20.9 (7.0)	20.4 (6.4)	19.2 (6.6)	16.9 (6.6)	15.1 (5.8)
Craft Story 21 recall delay. paraph., total units	16.3 (4.4)	15.9 (4.0)	15.2 (4.1)	13.6 (4.5)	12.1 (4.3)
Benson complex figure copy, total score	15.7 (1.2)	15.8 (1.2)	15.6 (1.3)	15.4 (1.6)	15.5 (1.5)
Benson complex figure recall, total score	12.6 (2.4)	11.9 (2.8)	11.2 (3.0)	10.1 (3.3)	9.4 (3.6)
Number span test forward, total correct trials	8.7 (2.4)	8.4 (2.3)	8.3 (2.3)	8.0 (2.3)	7.6 (2.1)
Number span test forward, longest span	6.9 (1.3)	6.7 (1.3)	6.7 (1.3)	6.5 (1.3)	6.3 (1.3)
Number span test backward, total correct trials	7.8 (2.3)	7.4 (2.2)	7.1 (2.2)	6.8 (2.2)	6.4 (2.4)
Number span test backward, longest span	5.4 (1.3)	5.2 (1.3)	5.1 (1.3)	4.9 (1.3)	4.7 (1.4)
MINT, total score	30.1 (2.0)	30.3 (2.0)	30.1 (2.2)	29.6 (2.6)	28.4 (3.4)
Phonemic test, F-words total in 60 s	16.3 (4.4)	15.5 (4.6)	14.9 (4.7)	14.6 (4.8)	13.6 (5.0)
Phonemic test, L-words total in 60 s	15.3 (4.3)	14.8 (4.4)	14.0 (4.4)	13.7 (4.5)	12.5 (4.7)
Phonemic test, total F-words and L-words	31.3 (8.3)	30.0 (8.4)	28.8 (8.5)	28.2 (8.6)	26.1 (9.1)
Animals list generation, total in 60 s	23.6 (5.3)	22.7 (5.5)	21.2 (5.4)	19.5 (5.6)	17.0 (5.4)
Vegetables list generation, total in 60 s	16.3 (4.1)	16.0 (4.3)	15.0 (4.2)	14.0 (4.3)	12.4 (4.1)
Trail making test part A, time (s)	22.3 (8.4)	27.3 (9.6)	31.0 (10.8)	36.2 (13.3)	42.0 (14.3)
Trail making test part B, time (s)	55.4 (27.2)	70.1 (33.8)	82.7 (41.4)	102.0 (55.2)	140.6 (75.3)

*Higher scores indicate better scores except for the trail making test parts A and B; higher scores indicate slower time to completion.
Delay. indicates delayed; Immed, immediate; MINT, multilingual naming test; Paraph, paraphrase; UDS, Uniform Data Set; Verbat, verbatim.

average squared differences between the observed score and the predicted score, which we substitute as an estimate for a population SD. The adjusted Z-scores are calculated for each test adjusting for a single demographic characteristic (ie, sex, age, or education) and adjusting for all 3 of these demographics. One can enter an individual's demographics and raw test scores, and the calculator uses the resulting Z-scores to calculate percentile estimates that indicate the individual's level of impairment on any given test (eg, low average, or severely impaired). Two new variables were also added to this calculator to improve the precision of percentile estimates for trail making part A and part B. These 2 tests are terminated if the subjects cannot complete within a specified time length (150 s and 300 s for A and B, respectively), resulting in the same score regardless of how many lines are correctly connected. We added connections-per-second (correct lines connected divided by the time to completion) for both Part A and B. These 2 new variables provide more accurate Z-scores and percentiles for the trail making tests.

TABLE 4. Mean Neuropsychological Test Scores by Education Group

UDS Version 3 Neuropsychological Test*	Mean (SD)			
	≤ 12 y	13-15 y	16 y	17+ y
MoCA, total score	24.1 (3.7)	25.7 (2.9)	26.6 (2.4)	26.9 (2.2)
Craft Story 21 recall immed. verbatim, total units	19.6 (7.1)	21.2 (6.5)	22.0 (6.5)	22.8 (6.4)
Craft Story 21 recall immed. paraphrase, total units	14.5 (4.7)	15.6 (4.0)	16.3 (4.1)	16.7 (3.8)
Craft Story 21 recall delay. verbatim, total units	16.6 (7.1)	18.3 (6.8)	19.2 (6.6)	20.0 (6.4)
Craft Story 21 recall delay. paraphrase, total units	13.2 (5.0)	14.4 (4.4)	15.3 (4.3)	15.8 (4.0)
Benson complex figure copy, total score	15.2 (1.5)	15.5 (1.4)	15.7 (1.3)	15.8 (1.2)
Benson complex figure recall, total score	10.2 (3.6)	11.0 (3.2)	11.4 (3.1)	11.5 (2.9)
Number span test forward, total correct trials	7.4 (2.2)	7.9 (2.2)	8.4 (2.2)	8.6 (2.3)
Number span test forward, longest span	6.2 (1.3)	6.5 (1.3)	6.8 (1.3)	6.9 (1.3)
Number span test backward, total correct trials	6.1 (2.3)	6.7 (2.2)	7.3 (2.2)	7.5 (2.2)
Number span test backward, longest span	4.5 (1.4)	4.9 (1.3)	5.2 (1.3)	5.3 (1.3)
MINT, total score	28.5 (2.9)	29.4 (2.4)	30.2 (2.0)	30.5 (2.0)
Phonemic test, F-words total in 60 s	12.8 (4.7)	13.9 (4.4)	15.2 (4.6)	16.1 (4.5)
Phonemic test, L-words total in 60 s	11.9 (4.5)	13.0 (4.4)	14.3 (4.1)	15.2 (4.3)
Phonemic test, total F-words and L-words	24.7 (8.6)	26.7 (8.3)	29.4 (8.1)	31.2 (8.2)
Animals list generation, total in 60 s	18.5 (5.4)	19.9 (5.3)	21.3 (5.1)	22.8 (5.7)
Vegetables list generation, total in 60 s	13.5 (4.0)	14.7 (4.0)	15.1 (4.1)	15.7 (4.5)
Trail making test part A, time (s)	35.1 (14.4)	32.1 (12.3)	29.7 (11.0)	29.2 (10.9)
Trail making test part B, time (s)	112.7 (69.8)	91.4 (51.5)	76.1 (36.4)	73.9 (35.9)

*Higher scores indicate better scores except for the trail making test parts A and B.
Delay. indicates delayed; Immed, immediate; MINT, multilingual naming test; MoCA, Montreal Cognitive Assessment; UDS, Uniform Data Set.

TABLE 5. Multivariable Linear Regression Coefficients and 95% CIs for Sex, Age, and Education

UDS Version 3 Neuropsychological Test[a]	Female	Age (y)	Education (y)
	Coefficient (95% CI)	Coefficient (95% CI)	Coefficient (95% CI)
MoCA, total score	**0.35** (0.18, 0.52)**	**−0.08** (−0.08, −0.07)**	**0.33** (0.30, 0.36)**
Craft Story 21 recall immed. verbatim	**0.96** (0.51, 1.40)**	**−0.13** (−0.15, −0.11)**	**0.41** (0.33, 0.48)**
Craft Story 21 recall immed. paraphrase	**0.60** (0.33, 0.88)**	**−0.08** (−0.09, −0.07)**	**0.28** (0.23, 0.33)**
Craft Story 21 recall delayed, verbatim	**0.78** (0.33, 1.23)**	**−0.15** (−0.17, −0.13)**	**0.43** (0.35, 0.51)**
Craft Story 21 recall delayed, paraphrase	**0.52** (0.23, 0.81)**	**−0.10** (−0.11, −0.09)**	**0.32** (0.27, 0.37)**
Benson complex figure copy, total score	0.04 (−0.05, 0.14)	**−0.01** (−0.02, −0.01)**	**0.07** (0.06, 0.09)**
Benson complex figure recall, total score	**−0.47** (−0.67, −0.26)**	**−0.09** (−0.09, −0.08)**	**0.16** (0.12, 0.19)**
Number span forward, total correct trials	**−0.29** (−0.45, −0.14)**	**−0.03** (−0.03, −0.02)**	**0.15** (0.13, 0.18)**
Number span forward, longest span	**−0.16** (−0.25, −.07)**	**−0.01** (−0.02, −0.01)**	**0.08** (0.07, 0.10)**
Number span backward, total correct trials	−0.10 (−0.25, 0.05)	**−0.03** (−0.04, −0.03)**	**0.17** (0.14, 0.20)**
Number span backward, longest span	−0.06 (−0.15, 0.03)	**−0.02** (−0.02, −0.01)**	**0.10** (0.08, 0.11)**
Multilingual naming test, total score	**−0.81** (−0.96, −0.66)**	**−0.03** (−0.04, −0.02)**	**0.22** (0.19, 0.25)**
Verbal fluency phonemic test, total correct F-words	**0.54** (0.22, 0.85)**	**−0.05** (−0.07, −0.04)**	**0.42** (0.36, 0.47)**
Verbal fluency phonemic test, total correct L-words	**0.63** (0.33, 0.92)**	**−0.05** (−0.07, −0.04)**	**0.45** (0.40, 0.50)**
Verbal fluency phonemic test, total correct F-words and L-words	**1.12** (0.55, 1.69)**	**−0.09** (−0.12, −0.07)**	**0.86** (0.76, 0.96)**
Category fluency: animals, total score	0.35 (−0.02, 0.71)	**−0.15** (−0.16, −0.13)**	**0.57** (0.50, 0.63)**
Category fluency: vegetables, total score	**2.49** (2.22, 2.77)**	**−0.08** (−0.10, −0.07)**	**0.31** (0.26, 0.36)**
Trail making test part A, time (s)	0.03 (−0.72, 0.78)	**0.45** (0.41, 0.48)**	**−0.73** (−0.86, −0.60)**
Trail making test part A, correct lines/time (s)	−0.00 (−0.02, 0.02)	**−0.01** (−0.01, −0.01)**	**0.02** (0.01, 0.02)**
Trail making test part B, time (s)	1.58 (−1.29, 4.45)	**1.64** (1.50, 1.77)**	**−4.65** (−5.15, −4.15)**
Trail making test part B, correct lines/time (s)	0.00 (−0.01, 0.01)	**−0.01** (−0.01, −0.01)**	**0.01** (0.01, 0.01)**

[a]Higher scores indicate better scores except for the trail making test parts A and B time in seconds.
**(Bold values) Statistically significant at $P < 0.01$.
CI indicates confidence interval; Immed, immediate; MoCA, Montreal Cognitive Assessment; UDS, Uniform Data Set.

DISCUSSION

This paper reports the results from a study to develop normative data for the version 3.0 revision of the Uniform Data Set Neuropsychological Battery. The complete UDS contains not only the neuropsychological battery but also demographic, medical, family history, neurological, biomarker, psychiatric, and functional data and available post mortem diagnosis on Clinical Core participants who have been followed longitudinally. Earlier versions have been collected since 2005 and stored in the database of the NACC at the University of Washington. All the data are available for sharing with researchers and therefore provide a rich source for generating hypotheses and investigating cognitive aging and dementia in a well-defined cohort.

The current revision of the neuropsychological battery provides an updated set of tests, targeting predominantly the symptoms of the most typical, amnestic, presentation of AD. The tests are nonproprietary and have the potential to increase sensitivity over former measures to very early symptoms of cognitive decline in older individuals with different levels of education. The new measures are similar to the old measures but have also enriched the standard data collection with novel scores to enhance available data using a relatively brief battery. The normative calculator provides a convenient tool to characterize the level of performance on the measures of the UDSNB 3.0 battery. The calculator and the battery are available online (www.alz.washington.edu/WEB/npsych_means.html), (www.alz.washington.edu/NONMEMBER/UDS/DOCS/VER3/UDS3_npsych_worksheets_C2.pdf).

There are some limitations to the study reported above. Although the ADCs encourage the participation of a diverse sample with respect to sex, education, and race, there was an overrepresentation of individuals who were white, female, and highly educated. Therefore, the findings are most relevant to research settings where these demographics are representative of the research volunteers. It will be important to expand the normative data for underrepresented groups and also for population-based samples. Another limitation is that the battery focuses on the spectrum from healthy cognition to dementia of the Alzheimer type and does not explicitly target symptoms associated with other forms of dementia. The CTF has introduced additional data collection modules, however, including specialized tests of symptoms related to frontotemporal dementia. Plans are under way to further expand clinical symptom assessment in other dementia syndromes. The availability of the UDSNB 3.0 at no cost to researchers will aid in encouraging more consistent and systematic data collection in disparate studies of cognitive aging and dementia.

REFERENCES

1. Beekly DL, Ramos EM, Lee WW, et al. The National Alzheimer's Coordinating Center (NACC) database: the uniform data set. *Alzheimer Dis Assoc Disord.* 2007;21:249–258.

2. Weintraub S, Salmon D, Mercaldo N, et al. The Alzheimer's Disease Centers' Uniform Data Set (UDS): the neuropsychologic test battery. *Alzheimer Dis Assoc Disord.* 2009;23:91–101.

3. Morris JC. The clinical dementia rating (CDR): current version and scoring rules. *Neurology.* 1993;43:2412–2414.

4. Morris JC, Weintraub S, Chui HC, et al. The uniform data set (UDS): clinical and cognitive variables and descriptive data from Alzheimer Disease Centers. *Alzheimer Dis Assoc Disord.* 2006;20:210–216.

5. Beekly DL, Ramos EM, van Belle G, et al. The National Alzheimer's Coordinating Center (NACC) Database: an Alzheimer disease database. *Alzheimer Dis Assoc Disord.* 2004;18:270–277.

6. Shirk SD, Mitchell MB, Shaughnessy LW, et al. A web-based normative calculator for the uniform data set (UDS) neuropsychological test battery. *Alzheimers Res Ther.* 2011;3:32.

7. Jack CR Jr, Knopman DS, Jagust WJ, et al. Hypothetical model of dynamic biomarkers of the Alzheimer's pathological cascade. *Lancet Neurol.* 2010;9:119–128.

8. Sperling RA, Aisen PS, Beckett LA, et al. Toward defining the preclinical stages of Alzheimer's disease: recommendations from the National Institute on aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimers Dement.* 2011;7:280–292.

9. Galasko D, Katzman R, Salmon DP, et al. Clinical and neuropathological findings in Lewy body dementias. *Brain Cogn.* 1996;31:166–175.

10. Salmon DP, Galasko D, Hansen LA, et al. Neuropsychological deficits associated with diffuse Lewy body disease. *Brain Cogn.* 1996;31:148–165.

11. Alladi S, Xuereb J, Bak T, et al. Focal cortical presentations of Alzheimer's disease. *Brain.* 2007;130 (pt 10):2636–2645.

12. Hof PR, Vogt BA, Bouras C, et al. Atypical form of Alzheimer's disease with prominent posterior cortical atrophy: a review of lesion distribution and circuit disconnection in cortical visual pathways. *Vision Res.* 1997;37:3609–3625.

13. Mendez MF, Ghajarania M, Perryman KM. Posterior cortical atrophy: clinical characteristics and differences compared to Alzheimer's disease. *Dement Geriatr Cogn Disord.* 2002;14:33–40.

14. Folstein MF, Folstein SE, McHugh PR. *Mini Mental State Examination.* Lutz, Florida: Psychological Assessment Resources; 2004.

15. Monsell SE, Dodge HH, Zhou XH, et al. Results from the NACC Uniform Data Set neuropsychological battery crosswalk study. *Alzheimer Dis Assoc Disord.* 2016;30:134–139.

16. Nasreddine ZS, Phillips N, Chertkow H. Normative data for the montreal cognitive assessment (MoCA) in a population-based sample. *Neurology.* 2012;78:765–766. author reply 766.

17. Nasreddine ZS, Phillips NA, Bedirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc.* 2005;53:695–699.

18. Wechsler D. *Wechsler Memory Scale-Revised Manual.* San Antonio, Texas: The Psychological Corporation; 1987.

19. Craft S, Newcomer J, Kanne S, et al. Memory improvement following induced hyperinsulinemia in Alzheimer's disease. *Neurobiol Aging.* 1996;17:123–130.

20. Ivanova I, Salmon DP, Gollan TH. The multilingual naming test in Alzheimer's disease: clues to the origin of naming impairments. *J Int Neuropsychol Soc.* 2013;19:272–283.

21. Larner AJ. Screening utility of the Montreal Cognitive Assessment (MoCA): in place of—or as well as—the MMSE? *Int Psychogeriatr.* 2012;24:391–396.

22. Trzepacz PT, Hochstetler H, Wang S, et al. Alzheimer's Disease Neuroimaging I. Relationship between the Montreal Cognitive Assessment and mini-mental state examination for assessment of mild cognitive impairment in older adults. *BMC Geriatr.* 2015;15:107.

23. Lam B, Middleton LE, Masellis M, et al. Criterion and convergent validity of the Montreal cognitive assessment with screening and standardized neuropsychological testing. *J Am Geriatr Soc.* 2013;61:2181–2185.

24. Goldstein FC, Ashley AV, Miller E, et al. Validity of the montreal cognitive assessment as a screen for mild cognitive impairment and dementia in African Americans. *J Geriatr Psychiatry Neurol.* 2014;27:199–203.

25. Rossetti HC, Lacritz LH, Cullum CM, et al. Normative data for the Montreal Cognitive Assessment (MoCA) in a population-based sample. *Neurology.* 2011;77:1272–1275.

26. Julayanont P, Tangwongchai S, Hemrungrojn S, et al. The Montreal Cognitive Assessment-basic: a screening tool for mild cognitive impairment in illiterate and low-educated elderly adults. *J Am Geriatr Soc.* 2015;63:2550–2554.

27. Luis CA, Keegan AP, Mullan M. Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the Southeastern US. *Int J Geriatr Psychiatry.* 2009;24:197–201.

28. Roalf DR, Moberg PJ, Xie SX, et al. Comparative accuracies of two common screening instruments for classification of Alzheimer's disease, mild cognitive impairment, and healthy aging. *Alzheimers Dement.* 2013;9:529–537.

29. Johnson N, Barion A, Rademaker A, et al. The activities of daily living questionnaire: a validation study in patients with dementia. *Alzheimer Dis Assoc Disord.* 2004;18:223–230.

30. Durant J, Leger GC, Banks SJ, et al. Relationship between the activities of daily living questionnaire and the montreal cognitive assessment. *Alzheimers Demen.* 2016;4:43–46.

31. Julayanont P, Brousseau M, Chertkow H, et al. Montreal Cognitive Assessment memory index score (MoCA-MIS) as a predictor of conversion from mild cognitive impairment to Alzheimer's disease. *J Am Geriatr Soc.* 2014;62:679–684.

32. Conti S, Bonazzi S, Laiacona M, et al. Montreal Cognitive Assessment (MoCA)-Italian version: regression based norms and equivalent scores. *Neurol Sci.* 2015;36:209–214.

33. McKhann G, Drachman D, Folstein M, et al. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer's disease. *Neurology.* 1984;34:939–944.

34. Craft S, Asthana S, Schellenberg G, et al. Insulin effects on glucose metabolism, memory, and plasma amyloid precursor protein in Alzheimer's disease differ according to apolipoprotein-E genotype. *Ann NY Acad Sci.* 2000;903:222–228.

35. Conroy SK, McDonald BC, Ahles TA, et al. Chemotherapy-induced amenorrhea: a prospective study of brain activation changes and neurocognitive correlates. *Brain Imaging Behav.* 2013;7:491–500.

36. Conroy SK, McDonald BC, Smith DJ, et al. Alterations in brain structure and function in breast cancer survivors: effect of post-chemotherapy interval and relation to oxidative DNA damage. *Breast Cancer Res Treat.* 2013;137:493–502.

37. Ferguson RJ, Ahles TA, Saykin AJ, et al. Cognitive-behavioral management of chemotherapy-related cognitive change. *Psychooncology.* 2007;16:772–777.

38. McDonald BC, Flashman LA, Arciniegas DB, et al. Methylphenidate and memory and attention adaptation training for persistent cognitive symptoms after traumatic brain injury: a randomized, placebo-controlled trial. *Neuropsychopharmacology.* 2017:1766–1775.

39. Possin KL, Laluz VR, Alcantar OZ, et al. Distinct neuroanatomical substrates and cognitive mechanisms of figure copy performance in Alzheimer's disease and behavioral variant frontotemporal dementia. *Neuropsychologia.* 2011;49:43–48.

40. Gollan TH, Weissburger G, Runnqvist E, et al. Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Biling : Lang Cogn.* 2011;13:215–218.

41. Possin KL. Visual spatial cognition in neurodegenerative disease. *Neurocase.* 2010;16:466–487.