




Research Article

Robust reference group normative data for neuropsychological tests accounting for primary language use in Asian American older adults

Arunima Kapoor¹, Jean K. Ho², Jung Yun Jang² and Daniel A. Nation^{3,4} 

¹Department of Psychological Science, University of California, Irvine, CA, USA, ²Institute for Memory Disorders and Neurological Impairments, University of California, Irvine, CA, USA, ³Leonard Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA and ⁴Department of Physiology and Neuroscience, Zilkha Neurogenetic Institute, University of Southern California, Keck School of Medicine, Los Angeles, CA, USA

Abstract

Objective: The present study aimed to develop neuropsychological norms for older Asian Americans with English as a primary or secondary language, using data from the National Alzheimer's Coordinating Center (NACC). **Method:** A normative sample of Asian American participants was derived from the NACC database using robust criteria: participants were cognitively unimpaired at baseline (i.e., no MCI or dementia) and remained cognitively unimpaired at 1-year follow-up. Clinical and demographic characteristics were compared between Primary and Secondary English speakers using analyses of variance for continuous measures and chi-square tests for categorical variables. Linear regression models compared neuropsychological performance between the groups, adjusting for demographics (age, sex, and education). Regression models were developed for clinical application to compute demographically adjusted z-scores. **Results:** Secondary English speakers were younger than Primary English speakers ($p < .001$). There were significant differences between the groups on measures of mental status (Mini-Mental State Examination, $p = .002$), attention (Trail Making Test A, Digit Span Forward Total Score, $p < .001$), language (Boston Naming Test, Animal Fluency, Vegetable Fluency, $p < .001$), and executive function (Trail Making Test B, $p = .02$). **Conclusions:** Separate normative data are needed for Primary vs. Secondary English speakers from Asian American backgrounds. We provide normative data on older Asian Americans to enable clinicians to account for English use in the interpretation of neuropsychological assessment scores.

Keywords: Dementia; cognition; neuropsychology; racial groups; Asians; language

(Received 12 December 2022; final revision 5 October 2023; accepted 9 October 2023)

Introduction

There are race-related differences in cognitive performance in later life (Masel et al., 2010; Sloan & Wang, 2005; Zsembik & Peek, 2001). At times, these disparities are attenuated by accounting for demographic characteristics, such as educational attainment (Barnes & Yaffe, 2011), literacy (Manly et al., 2004), reading level (Byrd et al., 2005), socioeconomic status (Schwartz et al., 2004), or health-related factors (Mungas et al., 2009). Thus, researchers have argued that “race” is merely a proxy for such differences (Dotson et al., 2008; Sisco et al., 2015). In the United States, normative standards derived from Non-Hispanic White, English-speaking populations are sometimes applied to ethnically and linguistically diverse examinees. This has resulted in increased rates of diagnostic errors and low-test specificity for such individuals (Byrd and Rivera-Mindt, 2022).

With growing populations of racially and ethnically diverse individuals in the United States, it is a priority in neuropsychology to improve methods for ascertaining the diagnosis of neurocognitive

disorders and to reduce rates of over- and underdiagnosis of cognitive impairment. Development of demographically adjusted norms, which account for fundamental sociocultural factors affecting neuropsychological performance, will advance our ability to evaluate individuals across a range of racial and ethnic groups with improved accuracy. Neuropsychological assessments can further benefit from specific norms that take into account demographic factors beyond just race alone, as this only captures one facet of an individual and may serve as a proxy for many other disparate factors such as literacy, language use, and education-related factors such as quality of education and educational attainment. In addition, instrument and test bias including the use of Latin alphabet and culturally-biased terminology in stimulus material may influence performance on neuropsychological assessments (Barker-Collo, 2001; Fernández & Abe, 2018).

To this end, there has been a multiplicity of efforts to test new standardization samples consisting of individuals from diverse backgrounds. Some examples in the United States include Mayo's Older African-American Normative Studies MOANS;

Corresponding author: D. A. Nation; Email: danation@usc.edu

Cite this article: Kapoor A., Ho J.K., Jang J.Y., & Nation D.A. Robust reference group normative data for neuropsychological tests accounting for primary language use in Asian American older adults. *Journal of the International Neuropsychological Society*, 1–8, <https://doi.org/10.1017/S1355617723000759>

© The Author(s), 2024. Published by Cambridge University Press on behalf of International Neuropsychological Society. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

(Lucas et al., 2005) and the Neuropsychological Norms for the US-Mexico Border Region in Spanish (NP-NUMBRs) Project (Rivera Mindt et al., 2021). Despite these efforts, there remains a dearth of normative data on many racial and ethnic groups, including Asian Americans in particular. Specifically, there is a dearth of robust norms for Asian Americans. Conventional norms are based on individuals studied at a single timepoint (De Santi et al., 2008; Holtzer et al., 2008); however, one limitation of conventional norming is that normative samples may include individuals who are in the preclinical stages of dementia and perform in the normal range on neuropsychological test (Sliwinski et al., 1996). Conversely, robust norming utilizes longitudinal assessment to exclude individuals who develop cognitive impairment at follow-up, thereby excluding individuals in the preclinical stages of disease (Holtzer et al., 2008; Koscik et al., 2014; Sliwinski et al., 2003). To date, few studies have established robust norms for neuropsychological tests in Asian American older adults.

In addition, one factor that has been scantily accounted for in the neuropsychological assessment of individuals from underrepresented and understudied groups is their primary language use. According to the 2019 US Census, the number of individuals who speak a language other than English at home has increased from 23.1 million in 1980 to 67.8 million in 2019, representing a 194% increase (Dietrich & Hernandez, 2022). Of these 67.8 million individuals, there were 3.49 million Chinese speakers, 1.76 million Tagalog speakers, 1.57 million Vietnamese speakers, and 1.08 million Korean speakers (Dietrich & Hernandez, 2022).

The present study sought to address the gap in available neuropsychological tools for Asian American assessment by providing normative neuropsychological data on Asian American individuals drawn from the National Alzheimer's Coordinating Center (NACC). Study aims included examination of whether participants' use of English – as a primary or secondary language – is an important factor to consider in normative practices. It was hypothesized that use of English as a primary vs. secondary language would be significantly related to neuropsychological performance. Thus, the study also aimed to create normative data accounting for the type of English use (i.e., primary vs. secondary).

Method

The current study was conducted in accordance with the World Medical Association Declaration of Helsinki.

Study population

This study involves secondary analysis of the National Alzheimer's Coordinating Center (NACC) database, obtained using the request form available at <https://www.naccdata.org/>. Data in the NACC database were from participants recruited at 33 Alzheimer's Disease Centers (ADCs) across the United States between September 2005 and February 2020. Participants underwent the same assessments and were evaluated for incident MCI and dementia at yearly intervals. Data from the first follow-up visits with these participants through February 2021 were included. The present study included 338 participants (Fig. 1) who met the following inclusion criteria at baseline: (1) were aged ≥ 55 years; (2) self-reported race as Asian or Asian American; (3) had at least one follow-up visit; (4) were diagnosed as cognitively healthy at baseline and at the first follow-up visit. We employed a robust norming approach, whereby all participants were cognitively healthy at least two timepoints (Holtzer et al., 2008). All contributing ADCs obtained informed consent from their

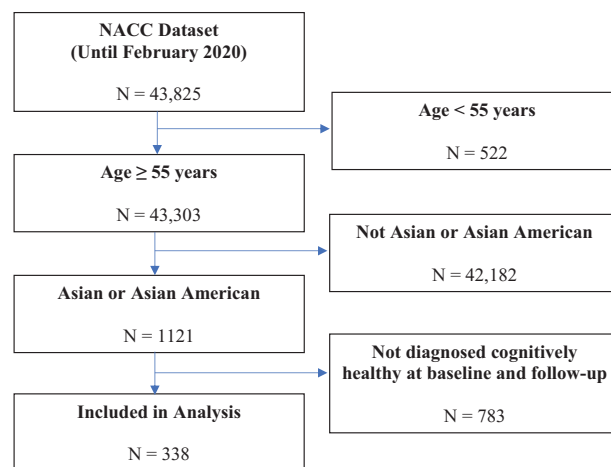


Figure 1. Study eligibility criteria.

participants and received approval from local institutional review boards.

Clinical diagnosis

Cognitive status was established based on neuropsychological testing and Clinical Dementia Rating (CDR) score, diagnosed by a single clinician or consensus panel as outlined in the NACC protocol. Normal cognition was established based on neuropsychological testing within normal range and/or global CDR score of 0. Independence in functional abilities, change in cognition, history and objective cognitive assessment were all considerations in determining diagnosis.

Neuropsychological tests

Neuropsychological tests were drawn from the Uniform Data Set versions 1, 2, and 3, and included the Mini-Mental State Exam (MMSE), Wechsler Memory Scale - Revised Logical Memory Story A Immediate Recall (Logical Memory I) and Delayed Recall (Logical Memory II), Wechsler Adult Intelligence Scale-Revised (WAIS-R) Digit Span Forward, Digit Span Backward, Animal Fluency, Vegetable Fluency, Trail Making Test Parts A and B, WAIS-R Digit Symbol, and the Boston Naming Test - 30-item version (BNT-30 odd-numbered items). Between the Uniform Data Set Version 2 and Version 3, the tests included in the battery were changed. We examined the data from all versions if the same tests were administered across all versions, and only examined tests from Version 1 and 2 if those tests were discontinued in Version 3. Higher scores indicate better performance on all tests except for the Trail Making Test, for which a higher score indicates longer time to completion and therefore worse performance.

Medical history

Body mass index (weight, height), systolic or diastolic blood pressures, and history of hypertension, diabetes, or depression was determined based on clinical evaluation during study visits. Stroke that affected cognition represents any history of stroke that had a relationship with cognitive impairment.

Statistical analysis

Clinical and demographic characteristics of the study sample were compared using t-tests, analyses of variance for continuous

Table 1. Demographics by primary language use

Ages		English	Mandarin	Cantonese	Japanese	Other	N (%)
5559	N	11	7	4	–	2	24 (8.0)
	Education (Years), M (SD)	16.27 (1.19)	17.00 (2.77)	15.50 (2.52)		16.50 (2.12)	
	Male (%)	4 (36.4)	1 (14.3)	2 (50.0)		0 (0.0)	
6064	N	19	13	3	2	6	
	Education (Years), M (SD)	17.26 (1.94)	17.31 (2.69)	10.00 (5.29)	18.50 (3.54)	15.33 (2.73)	43 (14.3)
	Male (%)	5 (26.3)	2 (15.4)	1 (33.3)	1 (50.0)	0 (0.00)	
6569	N	31	15	4	2	9	61 (20.3)
	Education (Years), M (SD)	17.39 (2.36)	16.07 (2.82)	17.75 (2.87)	17.00 (1.41)	16.89 (3.02)	
	Male (%)	15 (48.4)	5 (33.3)	1 (25.0)	0 (0.0)	5 (55.6)	
7074	N	43	12	5	2	8	70 (23.4)
	Education (Years), M (SD)	16.26 (3.10)	17.17 (2.21)	14.00 (2.55)	17.00 (1.41)	18.13 (1.36)	
	Male (%)	18 (41.9)	5 (41.7)	0 (0.0)	0 (0.00)	4 (50.0)	
7579	N	38	6	4	2	7	57 (19.0)
	Education (Years), M (SD)	17.11 (2.23)	19.00 (1.10)	13.50 (5.69)	13.00 (7.07)	19.14 (1.07)	
	Male (%)	19 (50.0)	1 (16.7)	0 (0.0)	0 (0.0)	4 (57.1)	
8084	N	28	3	–	–	–	34 (11.3)
	Education (Years), M (SD)	16.21 (2.04)	16.00 (4.00)				
	Male (%)	7 (25.0)	2 (66.7)				
85+	N	9	–	–	–	2	11 (3.7)
	Education (Years), M (SD)	15.00 (3.32)				20.00 (0.00)	
	Male (%)	4 (44.4)				2 (100.00)	
Total	N	179 (59.6)	56 (18.7)	21 (7.0)	9 (3.0)	35 (11.7)	300 (100)
	Education (Years), M (SD)	16.67 (2.49)	17.02 (2.62)	14.62 (4.27)	16.78 (3.73)	17.74 (2.75)	
	Male (%)	72 (40.2)	16 (28.6)	4 (19.0)	2 (22.2)	16 (45.7)	

Data is not presented for cells with $N = 1$.

measures and chi-square tests for categorical variables. Where applicable, we also conducted the same comparisons using non-parametric tests such as Mann-Whitney. To examine differences in baseline neuropsychological performance between the participants who used English as a primary language (“Primary English speakers”) and those who used it as a secondary language (“Secondary English speakers”), analyses of covariance was applied, adjusting for age, sex and education. We utilized partial eta squared (η^2p) as a measure of effect size, where 0.01 represents a small effect, 0.06 represents a medium effect and 0.14 represents a large effect size. Mean, standard deviation, median and inter-quartile range for each cognitive test for Primary and Secondary English speakers were also computed to illustrate differences between the two groups. These analyses were done to determine the need, if any, for separate normative data based on English use.

Furthermore, using baseline test scores, multiple regression equations were developed to estimate the effect of English use (0 = secondary, 1 = primary), age (in years), sex (0 = female, 1 = male), and education (in years) for each neuropsychological test in NACC separately. These equations can be used to obtain demographically adjusted z-scores and corresponding percentiles for tests commonly used in the diagnosis of dementia (Clark et al., 2016; De Santi et al., 2008; Shirk et al., 2011a). For any participant i ,

$$Y'_{\text{Test Score}} = \beta_{0j} + \beta_1 \text{PrimaryEnglishUse}_i + \beta_2 \text{Age}_i + \beta_3 \text{Sex}_i + \beta_4 \text{Education}_i \quad (1)$$

where Y' = the predicted population mean score for any one test, β_{0j} = random intercept for each test, and β_1 , β_2 , β_3 and β_4 = coefficients corresponding to English use, age, sex, and education. Obtaining z-scores for individual participants will follow the formula:

$$z = \frac{Y - Y'}{\text{RMSE}} \quad (2)$$

where z = z-score estimate for any one individual's performance on a neuropsychological test, Y = the raw score obtained by this individual on the test, Y' = the predicted population mean score, derived from Equation 1 detailed above, and RMSE = root mean square error of the regression equation. The RMSE is the square root of the average squared differences between observed and predicted scores. We also evaluated multicollinearity, normal P-P plot and scatterplots of the residuals for each model to evaluate assumptions. All analyses were performed using R version 3.6.2 software and SPSS for Mac OS X version 21.0 (SPSS, Armonk, NY: IBM Corp.).

Results

Clinical and demographic information

Table 1 shows primary language use frequencies in the sample as well as years of education and proportion of males and females among different age categories for each language. Table 2 shows the baseline clinical and demographic characteristics for the NACC-derived Asian American robustly normative sample (i.e., cognitively healthy at baseline and at 1-year follow-up). Secondary English speakers were younger than Primary English speakers ($p < .001$). However, there were no significant differences in years of education, sex, body mass index scores, systolic or diastolic blood pressures, global CDR scores, or proportions with hypertension, diabetes, or depression within the last 2 years (all p 's $> .05$). The average length of follow-up was 1.25 years ($SD = 0.50$). Within our sample, only 2 participants identified as having Hispanic/Latino ethnicity. “Other” languages (Table 1) included languages such as Vietnamese, Thai, Tagalog, and

Table 2. Baseline characteristics of the robustly normative subsample (those with at least 1 follow-up, and who are cognitively healthy at baseline and at 1-year follow-up)

		<i>N</i> (%) ¹	Primary English speakers (<i>N</i> = 198)	<i>N</i> (%) ²	Secondary English speakers (<i>N</i> = 140)	<i>p</i> -value	Cohen's <i>d</i> Cramer's <i>V</i>	Cliff's delta
Age (Years)	mean (<i>SD</i>) median (IQR)	181 (91.4)	72.73 (7.85) 73.00 (12.00)	123 (87.9)	68.58 (7.10) 68.00 (11.00)	< .001	-.55	0.70
Sex (Male)	<i>N</i> (%)	198 (100)	79 (39.9)	140 (100)	48 (34.3)	.29	.06	–
Education (Years)	mean (<i>SD</i>) median (IQR)	196 (99.0)	16.53 (2.65) 16.00 (2.00)	138 (98.6)	16.80 (3.11) 17.50 (2.00)	.40	.09	0.92
BMI	mean (<i>SD</i>)	171 (86.4)	24.39 (4.08)	119 (85.0)	23.51 (3.81)	.07	– .22	0.91
Systolic blood pressure	mean (<i>SD</i>)	170 (85.9)	132.64 (17.94)	117 (83.6)	134.02 (17.53)	.52	.08	0.97
Diastolic blood pressure	mean (<i>SD</i>)	170 (85.9)	75.53 (10.61)	116 (82.9)	76.11 (9.08)	.63	.06	0.94
CDR - Global Score	mean (<i>SD</i>) median (IQR)	181 (91.4)	0.07 (0.17) 0.00 (0.00)	123 (87.9)	0.06 (0.16) 0.00 (0.00)	.68	–.05	0.98
Stroke that affected cognition	<i>N</i> (%)	181 (91.4)	0 (0.0)	123 (87.9)	0 (0.0)	–	–	–
Hypertension	<i>N</i> (%)	180 (90.9)	89 (49.4)	121 (86.4)	46 (38.0)	.05	.11	–
Diabetes	<i>N</i> (%)	178 (89.9)	26 (14.6)	119 (85.0)	13 (10.9)	.36	.05	–
Depression within last 2 years	<i>N</i> (%)	180 (90.9)	18 (10.0)	121 (86.4)	7 (5.8)	.19	.08	–

BMI = Body Mass Index; CDR = Clinical Dementia Rating scale; SD = Standard Deviation; IQR = Interquartile Range.

Note: *T*-test and chi-square test were utilized for continuous and categorical variables, respectively. Cohen's *d* and Cramer's *V* were utilized for continuous and categorical variables, respectively.

¹Represents percentage of primary English speakers included in the analysis (*N* = 198).

²Represents percentage of secondary English speakers included in the analysis (*N* = 140).

Same results were obtained when we utilized non-parametric tests (i.e., Mann-Whitney).

Table 3. Differences on baseline neuropsychological performance between primary and secondary English speakers in the robust sample

Neuropsychological Test	<i>N</i>	Primary English speakers	<i>N</i>	Secondary English speakers	<i>p</i> -value	Effect size (η_p^2)
MMSE total score	118	28.82 (1.45)	57	28.35 (1.56)	.002	.054
LM I total score	117	13.21 (3.74)	57	13.19 (4.59)	.20	.009
LM II total score	117	12.25 (3.72)	57	12.40 (4.49)	.31	.006
DS Forward total score	117	7.82 (1.86)	58	6.72 (1.86)	<.001	.095
DS Forward span	117	6.21 (1.02)	58	5.83 (1.08)	.001	.062
DS Backward total score	117	6.62 (2.07)	58	6.52 (1.81)	.09	.017
DS Backward span	117	4.77 (1.09)	58	4.71 (1.03)	.07	.020
Animals total score	172	19.51 (5.36)	87	16.68 (4.04)	<.001	.098
Vegetables total score	172	14.54 (4.38)	87	12.98 (4.16)	<.001	.049
TMT-A total time	172	31.55 (10.61)	87	37.23 (17.59)	<.001	.066
TMT-B total time	171	84.85 (34.26)	86	87.26 (39.10)	.02	.020
Digit Symbol total score	115	50.70 (11.39)	56	52.07 (12.53)	.06	.021
BNT-30 total score	116	26.62 (3.98)	55	21.67 (5.57)	<.001	.199

MMSE = Mini-Mental State Exam; LM I = Wechsler Memory Scale - Revised (WMS-R) Logical Memory Story A Immediate Recall; LM II = WMS-R Logical Memory Story A Delayed Recall; DS = Digit Span; Animals = Animal Fluency; Vegetables = Vegetable Fluency; TMT-A = Trail Making Test Part A; TMT-B = Trail Making Test Part B; Digit Symbol = Wechsler Adult Intelligence Scaled - Revised (WAIS-R) Digit Symbol; BNT-30 = Boston Naming Test 30-item version.

Data are presented as Mean (SD). Higher scores indicate better performance for all tests except for Trail Making Test (Parts A and B) for which higher scores indicate longer time to completion and therefore worse performance. Significant differences between groups are indicated in boldface type.

F-test (ANCOVA) controlling for age, sex and education was utilized.

Korean. Testing was conducted in Mandarin or Cantonese for some participants (*N* = 35).

Neuropsychological performance

As shown in Table 3, there were significant group differences between Primary English speakers and Secondary English speakers on measures of memory, attention, and executive function, after correcting for covariates. Adjusting for age, sex and education, Secondary English speakers had significantly worse performance than the Primary English speakers on Trail Making Test B ($p = .02$), MMSE ($p = .002$), Digit Span Forward Total Score, Animal Fluency, Vegetable Fluency, Trail Making Test A, BNT-30 (all p 's < .001), and Digit Span Forward Span ($p = .001$).

To further illustrate these differences, summary statistics including mean, standard deviation, median and interquartile range

for Primary English and Secondary English speakers are presented in Table 4. Raw mean scores on all assessments at baseline and follow-up for primary and secondary English speakers are included in the Supplemental Materials (Supplemental Figures 1-13).

Table 5 presents the coefficients with 95% confidence intervals, as well as RMSE values, for our multivariate regression equations. The variance inflation factors were below 2 for all variables in every model. Based on evaluation of normal P-P plots and scatterplot of the residuals for each model, assumptions of linear regression were met and a linear model was deemed most appropriate.

Values from Table 5 can be used for estimating *z*-scores corresponding to various neuropsychological tests, accounting for English use, age, sex, and years of education. To illustrate the use of Table 5, the predicted mean BNT-30 score for a theoretical population of 70-year-old women with 12 years of education, who are Secondary English speakers (Primary English Use = 0,

Table 4. Baseline summary statistics for cognitively healthy participants

Neuropsychological test	Primary English speakers				Secondary English speakers			
	Max score	N	Mean (SD)	Median (IQR)	Max score	N	Mean (SD)	Median (IQR)
MMSE total score	30	119	28.81 (1.45)	29.00 (2.00)	30	57	28.35 (1.56)	29.00 (2.50)
LM I total score	21	118	13.22 (3.73)	13.00 (5.00)	24	57	13.19 (4.59)	14.00 (7.00)
LM II total score	20	118	12.26 (3.71)	12.00 (6.00)	24	57	12.40 (4.49)	13.00 (6.50)
DS Forward total score	12	118	7.84 (1.87)	8.00 (2.00)	12	58	6.72 (1.86)	7.00 (3.00)
DS Forward span	8	118	6.23 (1.03)	6.00 (1.00)	8	58	5.83 (1.08)	6.00 (2.00)
DS Backward total score	12	118	6.60 (2.06)	6.00 (3.00)	10	58	6.52 (1.81)	7.00 (3.00)
DS Backward span	7	118	4.76 (1.09)	5.00 (1.00)	7	58	4.71 (1.03)	5.00 (1.00)
Animals total score	35	174	19.54 (5.35)	19.00 (7.00)	29	89	16.72 (4.04)	16.00 (5.00)
Vegetables total score	29	174	14.52 (4.36)	14.00 (6.00)	25	89	12.96 (4.13)	13.00 (5.00)
TMT-A total time	74	174	31.48 (10.56)	30.00 (14.00)	133	89	37.01 (17.45)	32.00 (13.00)
TMT-B total time	240	173	84.52 (34.20)	80.00 (36.00)	300	88	87.20 (38.74)	77.50 (29.00)
Digit Symbol total score	86	116	50.69 (11.34)	50.00 (15.00)	83	56	52.07 (12.53)	51.00 (17.50)
BNT-30 total score	30	117	26.60 (3.97)	28.00 (3.50)	30	55	21.67 (5.57)	22.00 (6.00)

MMSE = Mini-Mental State Exam; LM I = Wechsler Memory Scale - Revised (WMS-R) Logical Memory Story A Immediate Recall; LM II = WMS-R Logical Memory Story A Delayed Recall; DS = Digit Span; Animals = Animal Fluency; Vegetables = Vegetable Fluency; TMT-A = Trail Making Test Part A; TMT-B = Trail Making Test Part B; Digit Symbol = Wechsler Adult Intelligence Scaled - Revised (WAIS-R) Digit Symbol; BNT-30 = Boston Naming Test 30-item version.

Table 5. Regression coefficients with 95% confidence intervals and the root mean square error (RMSE) for our multivariate regression equations, for estimating z – scores corresponding to various neuropsychological tests

Neuropsychological test	Intercept	Primary English speaker*	Age (y)	Male	Education (y)	RMSE
		b ₁ Coefficient (95% CI)	b ₂ Coefficient (95% CI)	b ₃ Coefficient (95% CI)	b ₄ Coefficient (95% CI)	
MMSE total score	27.85	0.75 (0.27, 1.22)	−0.03 (−0.06, −0.003)	−0.44 (−0.90, 0.02)	0.16 (0.07, 0.24)	1.415
LM I total score	12.88	0.82 (−0.46, 2.10)	−0.08 (−0.15, −0.01)	−2.03 (−3.27, −0.79)	0.39 (0.16, 0.61)	3.808
LM II total score	12.27	0.66 (−0.61, 1.92)	−0.09 (−0.16, −0.02)	−1.88 (−3.10, −0.65)	0.39 (0.16, 0.61)	3.768
DS Forward total score	8.23	1.32 (0.70, 1.94)	−0.04 (−0.07, −0.003)	0.17 (−0.43, 0.77)	0.06 (−0.05, 0.17)	1.843
DS Forward span	7.90	0.57 (0.23, 0.91)	−0.03 (−0.05, −0.02)	0.11 (−0.22, 0.43)	0.02 (−0.04, 0.08)	1.005
DS Backward total score	5.63	0.54 (−0.08, 1.15)	−0.05 (−0.08, −0.02)	−0.33 (−0.93, 0.27)	0.26 (0.15, 0.37)	1.836
DS Backward span	4.63	0.31 (−0.02, 0.65)	−0.03 (−0.05, −0.01)	−0.23 (−0.55, 0.10)	0.13 (0.07, 0.19)	0.995
Animals total score	18.55	3.44 (2.15, 4.72)	−0.11 (−0.19, −0.04)	0.26 (−1.02, 1.55)	0.34 (0.11, 0.58)	4.812
Vegetables total score	14.97	2.03 (0.92, 3.13)	−0.06 (−0.12, 0.01)	−2.59 (−3.69, −1.49)	0.16 (−0.04, 0.36)	4.125
TMT-A total time	7.24	−7.41 (−10.86, −3.96)	0.48 (0.27, 0.69)	−1.70 (−5.13, 1.74)	−0.16 (−0.79, 0.47)	12.892
TMT-B total time	16.39	−10.04 (−18.78, −1.31)	1.75 (1.22, 2.28)	−4.56 (−13.23, 4.10)	−2.81 (−4.41, −1.22)	32.510
Digit Symbol total score	80.50	3.29 (−0.19, 6.76)	−0.64 (−0.83, −0.45)	−4.62 (7.94, −1.30)	0.98 (0.35, 1.61)	10.153
BNT-30 total score	17.67	4.98 (3.45, 6.52)	0.01 (−0.08, 0.09)	1.12 (−0.35, 2.60)	0.19 (−0.09, 0.46)	4.511

LM II = WMS-R Logical Memory Story A Delayed Recall; DS = Digit Span; Animals = Animal Fluency; Vegetables = Vegetable Fluency; TMT-A = Trail Making Test Part A; TMT-B = Trail Making Test Part B; Digit Symbol = Wechsler Adult Intelligence Scaled - Revised (WAIS-R) Digit Symbol; BNT-30 = Boston Naming Test 30-item version.

MMSE = Mini-mental State Exam; LM I = Wechsler Memory Scale - Revised (WMS - R) Logical Memory Story A Immediate Recall;

*In Equation 1, Primary English Use = 1 for Primary English speakers, and Primary English Use = 0 for Secondary English Speakers.

Age = 70, Sex = 0, Education = 12), the following variables would be entered into Equation 1 to obtain a predicted BNT-30 total score of 20.65 out of 30 possible points.

$$Y'_{\text{BNT}} = 17.67 + (4.98 \times 0) + (0.01 \times 70) + (1.12 \times 0) + (0.19 \times 12) = 20.65 \quad (3)$$

To obtain a z-score corresponding to a BNT-30 score of 25 obtained by an individual who is a 70-year-old woman, with 12 years of education and who is a Secondary English speaker, we would then enter $Y'_{\text{BNT}} = 20.65$ and the RMSE score for MMSE from Table 5 into Equation 2:

$$z = \frac{25 - 20.65}{4.511} = 0.96$$

This z-score can then be looked up in any number of conversion tables for its corresponding percentile, i.e., 84%.

In contrast, if this same individual was scored using normative data developed using largely Non-Hispanic White, primary

English speakers (i.e., NACC norms), (Shirk et al., 2011b), they would receive a z-score of -0.724, i.e., 23%.

Excel files to calculate predicted and z-scores are included in the supplementary material. In addition, bootstrapped coefficients for each regression model are included in the Supplemental Tables (Supplemental Tables 1–13).

Discussion

This study presents normative data for older Asian American individuals using neuropsychological data from the NACC database, which to our knowledge, have not been published elsewhere. Our analysis included 338 individuals between the ages of 55 and 91 who identified as Asian or Asian American and were cognitively healthy at baseline and at first follow-up visit. Our analyses indicated significant neuropsychological differences among primary and secondary English speakers in a robustly normative sample, which consisted of older Asian Americans who were cognitively unimpaired at baseline and after 1-year follow-up. Differences between primary and secondary English language usage were observed on tests of mental status, attention, language

(verbal fluency and naming), and executive function, demonstrating the clear need for normative data to account for how English is used. Given the number of tests and cognitive domains that were influenced by type of English use (primary vs. secondary), regression equations were developed to account for English use, in addition to sex, age, and years of education. These equations may be used by clinicians and researchers who are assessing older Asian Americans to compute standardized scores (e.g., *z*-scores and percentile ranges) that are easily interpretable.

The regression equations provided by the present study may be of great value to the field. It is noteworthy that neuropsychological testing in older Asian Americans with English as a secondary language may activate multiple languages compared to primary English speakers. Research in bilingualism has elucidated two cognitive mechanisms that cause differences on neuropsychological performance between bilinguals and monolinguals. These mechanisms are (1) interference or competition between languages for use/selection, and (2) lower frequency of language-specific use, since each language is only spoken for some of the time (Rivera Mindt et al., 2008). These mechanisms may explain the robust bilingual disadvantages found on verbal tasks (Bialystok et al., 2008; Gollan et al., 2008; Gollan & Brown, 2006), even when tested solely in their dominant, first-acquired language (Gollan & Acenas, 2004; Ivanova & Costa, 2008). Research has largely shown that both languages in bilinguals are always active. The presence of consistent dual-language activation suggests that bilinguals need to exert a measure of inhibitory control while interacting with/in, and responding to only one language (Green, 1998).

Despite the more taxing cognitive processing that is necessitated, the longer amount of time taken is likely spuriously misinterpreted as slower and therefore poorer performance. The exception would be on measures of cognitive control, in which, unsurprisingly, bilinguals show subtle advantages (Bialystok & Martin, 2004; Bunge et al., 2002). It has been hypothesized that bilingualism may enhance domains such as executive function. However, this remains an area of active study and debate, given that others have argued that the bilingual advantage may not exist (Paap et al., 2015). Attitudes towards time and speed also vary across cultures and may influence performance on timed measures among individuals and cultures who do not prioritize speed or are not familiar with timed assessments (Agranovich et al., 2011). In the worst-case scenario, lower scores on tests among secondary English speakers may be inaccurately perceived as impaired. In other situations, clinicians may simply throw out lower scores that are otherwise uninterpretable given the lack of normative data in this population.

Prior studies have development norms for Mandarin-speaking and Spanish-speaking older adults (Qi et al., 2022; Stricks et al., 1998), however no prior study has development robust normative data accounting for primary language use in Asian American older adults. This study adds to the growing need for normative studies in secondary English speakers. Moreover, while prior norms were developed for specific languages or ethnic populations, the norms developed in this study included an adjustment for primary or secondary English use within a sample of Asian Americans, which may allow for more precise norms within this population.

One study limitation was the homogeneity in terms of years of education, as all our participants had a high school diploma or higher education, with the average participant for both Primary and Secondary English speakers having a college degree. Among individuals with fewer years of education, differences in

neuropsychological test scores between primary and secondary English speakers may be more pronounced and may be influenced by whether an individual attended an institution where instruction was in English. Additionally, Secondary English speakers reported the use of many different primary languages, including Mandarin (46%), Cantonese, (19%), Japanese (6%) and other languages (28%). These categories were necessarily collapsed into one ("Secondary English speakers") as cell numbers would be too small for analyses otherwise. Moreover, there were no data we could use to account for the degree of acculturation, where participants' main educational experiences took place, age at which a language was learned, level of proficiency and quality of education, which are important factors that may affect neuropsychological performance. This study also utilized self-report to determine primary language as opposed to a formal measure of language proficiency, which is a limitation. In addition, we could not determine practice effects at the follow-up visit. While practice effects may have resulted in improved performance at follow-up, given that cognitive status was determined based on clinical consensus using scores on numerous measures, it is unlikely that it affected diagnosis. Another limitation of this study was that robust norms were established based on normal cognition at two visits; however, future studies incorporating additional follow-up assessments could further enhance the robustness of these norms. It should also be acknowledged that the term "Asian American" can obfuscate the fact that this is a racially, culturally, and linguistically diverse group. Indeed, the term encompasses individuals with ethnic heritage from Asia (e.g., Chinese, Indians, Filipinos, Japanese, Koreans, Thai, Vietnamese, Cambodians, Hmong, Indonesians, Laotians, Pakistanis) as well as the Pacific Islands (i.e., Polynesia, Micronesia, and Melanesia). While subgroup analyses were underpowered in the current study, clinicians would do well to consider the unique history of individuals from any particular subgroup, as each was influenced differently by immigration policies, patterns, and experiences (Wong, 2000). Readers are also encouraged to review the excellent work by Riccio et al. (2014) and Wong and Fujii (2015) regarding crucial considerations and practical guidelines with regard to neuropsychological assessment of Asian Americans. Moreover, Ardila (2005) illustrates the cultural values underlying cognitive testing and highlights how factors such as the relationship and cultural differences between the examiner and examinee, test instruction interpretation, and the social situation of testing are all culture-dependent (Ardila, 2005). These factors may also play a role in influencing performance on neuropsychological assessments.

Another limitation is that this study only included individuals between the ages of 55 and 91, with education ranging from 6 to 25 years, speaking largely only 4 primary languages. Therefore, the findings of our study are likely most applicable to those represented in our sample. Moreover, these norms were developed based on a secondary analysis of a large dataset. While this allowed a large sample, the NACC database was not originally intended to be utilized for development of gold standard normative data. Accordingly, we were only able to create regression equations for tests with available data. In addition, cognitive status in this study was determined based on neuropsychological testing. It is possible that due to biases inherent in neuropsychological tests, non-English-speaking individuals may have been over- or under-identified as cognitively healthy. Moreover, the tests administered as part of NACC data collection were not available for certain language groups and testing was conducted in Mandarin or Cantonese for some participants. Therefore, there was some

variability in administration of tests for different language groups. Future studies are warranted to improve neuropsychological test stimuli, norms and diagnosis for non-English-speaking individuals and allow standardization of testing procedures in non-English-speaking populations.

Finally, it is important to acknowledge the limitations of race-based norms (Franzen et al., 2022). In this study, we aimed to account for primary language use to acknowledge differences among Asian Americans in language use. However, many neuropsychological measures are biased and may not be adequate for assessment in diverse populations. Screening tools for diverse population are available in numerous languages and can be administered to better capture cognitive functioning in different populations (Huang et al., 2018; Lim et al., 2021). Until additional research, training and novel instruments are available to enhance neuropsychological assessment for diverse populations, the adjusted norms may allow us to account for differences such as language use among Asian Americans. Moreover, robust norms for other cultural and ethnic populations are also lacking. Additional studies are warranted to develop culturally sensitive tests and robust norms.

Despite the limitations detailed above, the present study represents a significant advance for the field given the paucity of normative data available for older Asian Americans at risk for dementia. The present study benefits from additional strengths. First, the study utilized a robustly normal sample undergoing the NACC neuropsychological battery, which consists of many tests that target the most common presentations of age-related neurodegenerative conditions, including Alzheimer's disease. Second, this is the only study, to our knowledge, that takes into account English usage (as a primary vs. secondary language) in providing normative data for individuals from underrepresented backgrounds. The way in which English is used may be considered a proxy for other sociocultural factors that the present study was not able to evaluate, such as acculturation, noted above.

Further development of normative data for individuals from underrepresented backgrounds will improve our ability to determine a patient's cognitive status more accurately. This, in turn, will have important implications for neuropsychological research and clinical practice in underserved and understudied populations.

Supplementary material. The supplementary material for this article can be found at <https://doi.org/10.1017/S1355617723000759>.

Author contribution. Arunima Kapoor, Jean K. Ho, and Jung Yun Jang contributed equally.

Funding statement. Dr Nation is funded by the National Institute on Aging (R01AG064228; R01AG060049; R01AG082073; P01AG052350). The NACC database is funded by NIA/NIH Grant U24 AG072122. NACC data are contributed by the NIA-funded ADCs: P50 AG005131 (PI James Brewer, MD, PhD), P50 AG005133 (PI Oscar Lopez, MD), P50 AG005134 (PI Bradley Hyman, MD, PhD), P50 AG005136 (PI Thomas Grabowski, MD), P50 AG005138 (PI Mary Sano, PhD), P50 AG005142 (PI Helena Chui, MD), P50 AG005146 (PI Marilyn Albert, PhD), P50 AG005681 (PI John Morris, MD), P30 AG008017 (PI Jeffrey Kaye, MD), P30 AG008051 (PI Thomas Wisniewski, MD), P50 AG008702 (PI Scott Small, MD), P30 AG010124 (PI John Trojanowski, MD, PhD), P30 AG010129 (PI Charles DeCarli, MD), P30 AG010133 (PI Andrew Saykin, PsyD), P30 AG010161 (PI David Bennett, MD), P30 AG012300 (PI Roger Rosenberg, MD), P30 AG013846 (PI Neil Kowall, MD), P30 AG013854 (PI Robert Vassar, PhD), P50 AG016573 (PI Frank LaFerla, PhD), P50 AG016574 (PI Ronald Petersen, MD, PhD), P30 AG019610 (PI Eric Reiman, MD), P50 AG023501 (PI Bruce Miller, MD), P50 AG025688 (PI Allan Levey, MD, PhD), P30 AG028383 (PI Linda Van Eldik, PhD), P50

AG033514 (PI Sanjay Asthana, MD, FRCP), P30 AG035982 (PI Russell Swerdlow, MD), P50 AG047266 (PI Todd Golde, MD, PhD), P50 AG047270 (PI Stephen Strittmatter, MD, PhD), P50 AG047366 (PI Victor Henderson, MD, MS), P30 AG049638 (PI Suzanne Craft, PhD), P30 AG053760 (PI Henry Paulson, MD, PhD), P30 AG066546 (PI Sudha Seshadri, MD), P20 AG068024 (PI Erik Roberson, MD, PhD), P20 AG068053 (PI Marwan Sabbagh, MD), P20 AG068077 (PI Gary Rosenberg, MD), P20 AG068082 (PI Angela Jefferson, PhD), P30 AG072958 (PI Heather Whitson, MD), P30 AG072959 (PI James Leverenz, MD).

Competing interests. None.

References

- Agranovich, A. V., Panter, A. T., Puente, A. E., & Touradji, P. (2011). The culture of time in neuropsychological assessment: Exploring the effects of culture-specific time attitudes on timed test performance in Russian and American samples. *Journal of the International Neuropsychological Society*, 17(4), 692–701. <https://doi.org/10.1017/S1355617711000592>
- Ardila, A. (2005). Cultural values underlying psychometric cognitive testing. *Neuropsychology Review*, 15(4), 185. <https://doi.org/10.1007/s11065-005-9180-y>
- Barker-Collo, S. L. (2001). The 60-item Boston naming test: Cultural bias and possible adaptations for New Zealand. *Aphasiology*, 15(1), 85–92. <https://doi.org/10.1080/02687040042000124>
- Barnes, D. E., & Yaffe, K. (2011). The projected effect of risk factor reduction on Alzheimer's disease prevalence. *The Lancet Neurology*, 10(9), 819–828. [https://doi.org/10.1016/S1474-4422\(11\)70072-2](https://doi.org/10.1016/S1474-4422(11)70072-2)
- Bialystok, E., Craik, F., & Luk, G. (2008). Cognitive control and lexical access in younger and older bilinguals. *Journal of Experimental Psychology: Learning Memory and Cognition*, 34(4), 859–873. <https://doi.org/10.1037/0278-7393.34.4.859>
- Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, 7(3), 325–339. <https://doi.org/10.1111/j.1467-7687.2004.00351.x>
- Bunge, S. A., Dudukovic, N. M., Thomason, M. E., Vaidya, C. J., & Gabrieli, J. D. E. (2002). Immature frontal lobe contributions to cognitive control in children: Evidence from fMRI. *Neuron*, 33(2), 301–311. [https://doi.org/10.1016/S0896-6273\(01\)00583-9](https://doi.org/10.1016/S0896-6273(01)00583-9)
- Byrd, D. A., Jacobs, D. M., Hilton, H. J., Stern, Y., & Manly, J. J. (2005). Sources of errors on visuoperceptual tasks: Role of education, literacy, and search strategy. *Brain and Cognition*, 58(3), 251–257. <https://doi.org/10.1016/j.bandc.2004.12.003>
- Byrd, D. A., & Rivera-Mindt, M. G. (2022). Neuropsychology's race problem does not begin or end with demographically adjusted norms. *Nature Reviews Neurology*, 18(3), 125–126. <https://doi.org/10.1038/s41582-021-00607-4>
- Clark, L. R., Kosciak, R. L., Nicholas, C. R., Okonkwo, O. C., Engelman, C. D., Bratzke, L. C., Hogan, K. J., Mueller, K. D., Bendlin, B. B., Carlsson, C. M., Asthana, S., Sager, M. A., Hermann, B. P., & Johnson, S. C. (2016). Mild cognitive impairment in late middle age in the wisconsin registry for Alzheimer's Prevention study: Prevalence and characteristics using robust and standard neuropsychological normative data. *Archives of Clinical Neuropsychology*, 31(7), 675–688. <https://doi.org/10.1093/arclin/acw024>
- De Santi, S., Pirraglia, E., Barr, W., Babb, J., Williams, S., Rogers, K., Glodzik, L., Brys, M., Mosconi, L., Reisberg, B., Ferris, S., & de Leon, M. J. (2008). Robust and conventional neuropsychological norms: Diagnosis and prediction of age-related cognitive decline. *Neuropsychology*, 22(4), 469–484. <https://doi.org/10.1037/0894-4105.22.4.469>
- Dietrich, S., & Hernandez, E. (2022). *Language use in the United States: 2019*. United States Census Bureau. <https://www.census.gov/content/dam/Census/library/publications/2022/acs/acs-50.pdf>
- Dotson, V. M., Kitner-Triolo, M., Evans, M. K., & Zonderman, A. B. (2008). Literacy-based normative data for low socioeconomic status African Americans. *The Clinical Neuropsychologist*, 22(6), 989–1017. <https://doi.org/10.1080/13854040701679017>

- Franzen, S., Pomati, S., Papma, J. M., Nielsen, T. R., Narme, P., Mukadam, N., Lozano-Ruiz, Álvaro, Ibanez-Casas, I., Goudsmit, M., Fasfous, A., Daugherty, J. C., Canevelli, M., Calia, C., van den Berg, E., & Bekkhus-Wetterberg, P. (2022). Cross-cultural neuropsychological assessment in Europe: Position statement of the European consortium on cross-cultural neuropsychology (ECCroN). *The Clinical Neuropsychologist*, 36(3), 546–557. <https://doi.org/10.1080/13854046.2021.1981456>
- Fernández, A. L., & Abe, J. (2018). Bias in cross-cultural neuropsychological testing: Problems and possible solutions. *Culture and Brain*, 6(1), 1–35. <https://doi.org/10.1007/s40167-017-0050-2>
- Gollan, T. H., & Acenas, L. A. R. (2004). What is a TOT? Cognate and translation effects on tip-of-the-tongue states in Spanish-English and Tagalog-English Bilinguals. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(1), 246–269. <https://doi.org/10.1037/0278-7393.30.1.246>
- Gollan, T. H., & Brown, A. S. (2006). From tip-of-the-tongue (TOT) data to theoretical implications in two steps: When more TOTs means better retrieval. *Journal of Experimental Psychology: General*, 135(3), 462–483. <https://doi.org/10.1037/0096-3445.135.3.462>
- Gollan, T. H., Montoya, R. I., Cera, C., & Sandoval, T. C. (2008). More use almost always means a smaller frequency effect: Aging, bilingualism, and the weaker links hypothesis. *Journal of Memory and Language*, 58(3), 787–814. <https://doi.org/10.1016/j.jml.2007.07.001>
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and Cognition*, 1(2), 67–81. <https://doi.org/10.1017/S1366728998000133>
- Holtzer, R., Goldin, Y., Molly Zimmerman, M., Katz, M., Buschke, H., & Lipton, R. B. (2008). Robust norms for selected neuropsychological tests in older adults. *Archives of Clinical Neuropsychology*, 23(5), 531–541. <https://doi.org/10.1016/j.acn.2008.05.004>
- Huang, L., Chen, K.-L., Lin, B.-Y., Tang, L., Zhao, Q.-H., Lv, Y.-R., & Guo, Q.-H. (2018). Chinese version of Montreal cognitive assessment basic for discrimination among different severities of Alzheimer's disease. *Neuropsychiatric Disease and Treatment*, 14, 2133–2140. <https://doi.org/10.2147/NDT.S174293>
- Ivanova, I., & Costa, A. (2008). Does bilingualism hamper lexical access in speech production? *Acta Psychologica*, 127(2), 277–288. <https://doi.org/10.1016/j.actpsy.2007.06.003>
- Koscik, R. L., La Rue, A., Jonaitis, E. M., Okonkwo, O. C., Johnson, S. C., Bendlin, B. B., Hermann, B. P., & Sager, M. A. (2014). Emergence of mild cognitive impairment in late middle-aged adults in the Wisconsin registry for Alzheimer's prevention. *Dementia and Geriatric Cognitive Disorders*, 38(1–2), 16–30. <https://doi.org/10.1159/000355682>
- Lim, S., Chong, S., Min, D., Mohaimin, S., Roberts, T., Trinh-Shevrin, C., & Kwon, S. C. (2021). Alzheimer's disease screening tools for Asian Americans: A scoping review. *Journal of Applied Gerontology*, 40(10), 1389–1398. <https://doi.org/10.1177/0733464820967594>
- Lucas, J. A., Ivnik, R. J., Willis, F. B., Ferman, T. J., Smith, G. E., Parfitt, F. C., Petersen, R. C., & Graff-Radford, N. R. (2005). Mayo's older African Americans normative studies: Normative data for commonly used clinical neuropsychological measures. *The Clinical Neuropsychologist*, 19(2), 162–183. <https://doi.org/10.1080/13854040509045265>
- Manly, J. J., Byrd, D., Touradji, P., Sanchez, D., & Stern, Y. (2004). Literacy and cognitive change among ethnically diverse elders. *International Journal of Psychology*, 39(1), 47–60. <https://doi.org/10.1080/00207590344000286>
- Masel, M. C., Raji, M., & Peek, M. K. (2010). Education and physical activity mediate the relationship between ethnicity and cognitive function in late middle-aged adults. *Ethnicity and Health*, 15(3), 283–302. <https://doi.org/10.1080/13557851003681273>
- Mungas, D., Reed, B. R., Farias, S. T., & DeCarli, C. (2009). Age and education effects on relationships of cognitive test scores with brain structure in demographically diverse older persons. *Psychology and Aging*, 24(1), 116–128. <https://doi.org/10.1037/a0013421>
- Paap, K. R., Johnson, H. A., & Sawi, O. (2015). Bilingual advantages in executive functioning either do not exist or are restricted to very specific and undetermined circumstances. *Cortex*, 69, 265–278. <https://doi.org/10.1016/j.cortex.2015.04.014>
- Qi, W., Sun, X., & Hong, Y. (2022). Normative data for adult mandarin-speaking populations: A systematic review of performance-based neuropsychological instruments. *Journal of the International Neuropsychological Society*, 28(5), 520–540. <https://doi.org/10.1017/S1355617721000667>
- Riccio, C. A., Yoon, H., & McCormick, A. S. (2014). Neuropsychological test selection with clients who are Asian. In R. Davis, & J. D. Amato (Eds.), *Issues of diversity in clinical neuropsychology* (pp. 151–174). Springer.
- Rivera Mindt, M., Arentoft, A., Kubo Germano, K., D'Aquila, E., Scheiner, D., Pizzirusso, M., Sandoval, T. C., & Gollan, T. H. (2008). Neuropsychological, cognitive, and theoretical considerations for evaluation of bilingual individuals. *Neuropsychology Review*, C(), 255–268. <https://doi.org/10.1007/s11065-008-9069-7>
- Rivera Mindt, M., Marquine, M. J., Aghvinian, M., Paredes, A. M., Kamalyan, L., Suárez, P., Heaton, A., Scott, T. M., Gooding, A., Diaz-Santos, M., Umlauf, A., Taylor, M. J., Artioli i Fortuny, L., Heaton, R. K., & Cherner, M. (2021). The neuropsychological norms for the U.S.-Mexico border region in Spanish (NP-NUMBRS) project: Overview and considerations for life span research and evidence-based practice. *The Clinical Neuropsychologist*, 35(2), 466–480. <https://doi.org/10.1080/13854046.2020.1794046>
- Schwartz, B. S., Glass, T. A., Bolla, K. I., Stewart, W. F., Glass, G., Rasmussen, M., Bressler, J., Shi, W., & Bandeen-Roche, K. (2004). Disparities in cognitive functioning by race/ethnicity in the Baltimore memory study. *Environmental Health Perspectives*, 112(3), 314–320. <https://doi.org/10.1289/ehp.6727>
- Shirk, S. D., Mitchell, M. B., Shaughnessy, L. W., Sherman, J. C., Locascio, J. J., Weintraub, S., & Atri, A. (2011a). A web-based normative calculator for the uniform data set (UDS) neuropsychological test battery. *Alzheimer's Research & Therapy*, 3(6), 32. <https://doi.org/10.1186/alzrt94>
- Shirk, S. D., Mitchell, M. B., Shaughnessy, L. W., Sherman, J. C., Locascio, J. J., Weintraub, S., & Atri, A. (2011b). A web-based normative calculator for the uniform data set (UDS) neuropsychological test battery. *Alzheimer's Research and Therapy*, 3(6), 32. <https://doi.org/10.1186/alzrt94>
- Sisco, S., Gross, A. L., Shih, R. A., Sachs, B. C., Glymour, M. M., Bangen, K. J., Benitez, A., Skinner, J., Schneider, B. C., & Manly, J. J. (2015). The role of early-life educational quality and literacy in explaining racial disparities in cognition in late life. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, 70(4), 557–567. <https://doi.org/10.1093/geronb/gbt133>
- Sliwinski, M. J., Hofer, S. M., Hall, C., Buschke, H., & Lipton, R. B. (2003). Modeling memory decline in older adults: The importance of preclinical dementia, attrition, and chronological age. *Psychology and Aging*, 18(4), 658–671. <https://doi.org/10.1037/0882-7974.18.4.658>
- 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*, 51B(4), P217–P225. <https://doi.org/10.1093/geronb/51B.4.P217>
- Sloan, F. A., & Wang, J. (2005). Disparities among older adults in measures of cognitive function by race or ethnicity. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, 60(5), P242–P250. <https://doi.org/10.1093/geronb/60.5.P242>
- Stricks, L., Pittman, J., Jacobs, D. M., Sano, M., & Stern, Y. (1998). Normative data for a brief neuropsychological battery administered to English- and Spanish-speaking community-dwelling elders. *Journal of the International Neuropsychological Society*, 4(4), 311–318.
- Wong, T. M. (2000). Neuropsychological assessment and intervention with Asian Americans. In *Handbook of cross-cultural neuropsychology* (pp. 43–53). Springer.
- Wong, T. M., & Fujii, D. E. (2015). Neuropsychological assessment of Asian Americans: Demographic factors, cultural diversity, and practical guidelines. *Cultural Diversity: A Special Issue of Applied Neuropsychology*, 4282, 23–36. <https://doi.org/10.4324/9780203764497-4>
- Zsembik, B. A., & Peek, M. K. (2001). Race differences in cognitive functioning among older adults. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, 56(5), 266–274. <https://doi.org/10.1093/geronb/56.5.S266>