

Research Article

# Measuring Cognitive Health in Ethnically Diverse Older Adults

Hector Hernandez Saucedo, BS,<sup>1,\*</sup> Rachel A. Whitmer, PhD,<sup>2,3</sup> Maria Glymour, ScD,<sup>4</sup> Charles DeCarli, MD,<sup>1</sup> Elizabeth-Rose Mayeda, PhD, MPH,<sup>5</sup> Paola Gilsanz, ScD,<sup>3,\*</sup> Sunita Q. Miles, MPH,<sup>3</sup> Nihal Bhulani, MD,<sup>1</sup> Sarah Tomaszewski Farias, PhD,<sup>1</sup> John Olichney, MD,<sup>1</sup> and Dan Mungas, PhD<sup>1</sup>

<sup>1</sup>Department of Neurology, University of California Davis, USA. <sup>2</sup>Public Health Sciences, University of California Davis, USA. <sup>3</sup>Kaiser Permanente Division of Research, Oakland, California, USA. <sup>4</sup>Epidemiology and Biostatistics, University of California San Francisco, USA. <sup>5</sup>Fielding School of Public Health, University of California Los Angeles, USA.

\*Address correspondence to: Hector Hernandez Saucedo, BS, UCD Alzheimer's Disease Center, 4860 Y Street, Suite 3900, Sacramento, CA 95831, USA. E-mail: [hhernandez@ucdavis.edu](mailto:hhernandez@ucdavis.edu)

Received: September 1, 2020; Editorial Decision Date: March 26, 2021

**Decision Editor:** Vanessa Taler, PhD

## Abstract

**Objectives:** Understanding racial/ethnic disparities in late-life cognitive health is a public health imperative. We used baseline data from the Kaiser Healthy Aging and Diverse Life Experiences (KHANDLE) study to examine how age, education, gender, and clinical diagnosis, a proxy for brain health, are associated with cross-sectional measures of cognition in diverse racial/ethnic groups.

**Methods:** Comprehensive measures of cognition were obtained using the Spanish and English Neuropsychological Assessment Scales and the National Institutes of Health Toolbox Cognitive Health Battery in a sample of 1,695 KHANDLE participants (Asians 24%, Blacks 26%, Latinos 20%, Whites 29%). A 25% random subsample was clinically evaluated and diagnosed with normal cognition, mild cognitive impairment (MCI), or dementia. Cognitive test scores were regressed on core demographic variables and diagnosis in the combined sample and in multiple group analyses stratified by racial/ethnic group.

**Results:** Race/ethnicity and education were variably associated with test scores with strongest associations with tests of vocabulary and semantic memory. Older age was associated with poorer performance on all measures, and gender differences varied across cognitive tests. Clinical diagnosis of MCI or dementia was associated with average decrements in test scores that ranged from  $-0.41$  to  $-0.84$  *SD*, with largest differences on tests of executive function and episodic memory. With few exceptions, associations of demographic variables and clinical diagnosis did not differ across racial/ethnic groups.

**Discussion:** The robust associations of cognitive test results with clinical diagnosis independent of core demographic variables and race/ethnicity support the validity of cognitive tests as indicators for brain health in diverse older adults.

**Keywords:** Cognition, Cross-cultural differences, Epidemiology, Neuropsychology

The continuous growth in racial/ethnic diversity of the older adult population in the United States (Treas & Carreon, 2010) and the urgent public health burden of

cognitive decline and dementia (Cotter, 2007) make it critically important that we understand the underpinnings of cognitive health in this increasingly diverse population.

Growing evidence suggests that racial/ethnic minorities including Blacks and Latinos may bear the highest Alzheimer's disease (AD) and dementia burden through the year 2060 (Matthews et al., 2019). In contrast to these rapidly emerging demographic trends, much of what we currently know about cognitive aging and cognitive health is derived from studies that inadequately represent our population's diversity. Consequently, there is a compelling need for studies in diverse populations that characterize late-life cognitive health and cognitive decline to provide a foundation for understanding racial/ethnic similarities and differences and identify interventions to promote successful cognitive aging.

Cognitive tests play a central role in research and clinical care related to cognitive aging and cognitive health. Cognitive test results are used to identify and diagnose clinically relevant cognitive impairment and monitor progression of diseases of aging like AD and related dementias (ADRD). Neuropsychological tests have long been used as indicators of both cognitive ability and brain health, but precise measurement is especially challenging in the context of demographically diverse populations. Specifically, demographic variables like racial/ethnic group membership (Brewster et al., 2014; Cagney & Lauderdale, 2002; Castora-Binkley et al., 2015; Early et al., 2013; Gross et al., 2015; Masel & Peek, 2009; Weuve et al., 2018), educational exposure (Early et al., 2013; Masel & Peek, 2009; Wilson et al., 2016), and quality of education (Glymour & Manly 2008; Manly et al., 2002; Sisco et al., 2015) strongly influence test scores, with robust effects that are present even after accounting for evidence of brain injury. Cognitive performance at a single time point is complexly determined not only by brain health but also by factors that are independent of brain health including lifetime experiences and exposures (Brewster et al., 2014; Melrose et al., 2015), and familiarity with cognitive testing (Early et al., 2013; Karlamangla et al., 2009). Indeed, surrogate measures of brain health such as regional magnetic resonance imaging (MRI) atrophy or vascular injury account for less than 50% of the variance in cognitive tests (Dowling et al., 2011; Mungas et al., 2009).

Despite evidence that longitudinal change in cognitive function is a better, more specific indicator of brain health and can better predict the presence of neurodegenerative disorders (ADRD) or progressive cerebrovascular disease (Early et al., 2013; Fletcher et al., 2018; Mungas et al., 2010; Walter et al., 2019), longitudinal follow-up is often not available, especially in clinical settings. We propose that cross-sectional test performance can be used to characterize brain health if effects that are not due to brain health (e.g., cultural, educational, and linguistic factors) can be accurately parceled out. Concretely, the goal is to compare a given individual's cross-sectional test scores with what would be expected of cognitively normal persons who have the same demographic characteristics and

life exposures. But this requires empirical data to establish the effects of nonbrain variables within the specific populations of interest, and research to address this question requires diverse samples that represent the older adult population.

Understanding demographic influences on cognitive tests and how these effects differ by race/ethnicity is a prerequisite for parsing the effects of brain and disease variables on cross-sectional measures of cognition. There is evidence that effects of age (Byrd et al., 2018) and education (Díaz-Venegas et al., 2016; Jean et al., 2019) on cognitive test scores might differ across racial/ethnic groups. AD and related degenerative diseases are increasingly prevalent with advancing age and may underlie much of the cognitive decline associated with age and increased dementia risk (Fjell et al., 2014). Racial/ethnic group differences in age effects on cross-sectional cognitive test scores could alternately result from greater sensitivity of test scores to disease effects in some groups, rather than increased disease prevalence or severity. Differential effects of education could also have important implications for determining normative expectations for performance that are critical for diagnosing cognitive impairment using cross-sectional test results. Similarly, racial/ethnic differences in gender effects on cognitive test scores are relevant for understanding normative expectations.

In the current study, we examined how different cross-sectional measures of cognition are influenced by core demographic variables including education, age, and gender in four major racial/ethnic groups: Latinos, non-Latino Asians, non-Latino Blacks, and non-Latino Whites. We used baseline assessment data from the Kaiser Healthy Aging and Diverse Life Experiences (KHANDLE) study, a large cohort study of members in the Kaiser Permanente Northern California Health System (KPNC). We measured cognitive health using the Spanish and English Neuropsychological Assessment Scales (SENAS; Mungas et al., 2004) and the National Institutes of Health (NIH) Toolbox Cognitive Health Battery (NIHTB-CHB; Weintraub et al., 2013). These two rigorous assessment instruments provided comprehensive assessment of multiple cognitive domains including episodic memory, executive function, and semantic memory. Clinical diagnosis was available in a randomly selected subsample and was used as a proxy for brain health. The purpose of this study was to inform the use of cognitive tests in diverse populations to assess and monitor brain health by: (a) examining how core demographic variables relate to measures of specific cognitive domains and whether these associations differ across racial/ethnic groups, and (b) examining how test scores are associated with clinical diagnosis independent of demographic variables and whether these associations differ across groups. Our overarching hypothesis was that cross-sectional measures of cognition are effective indicators of brain health in racially/ethnically and demographically diverse older adults.

## Method

### Participants

We used baseline data from the KHANDLE cohort, comprised of community-dwelling older adults residing in Northern California in the San Francisco Bay area and Sacramento valley. KHANDLE aims to evaluate how life course and sociocultural factors influence late-life brain health and cognitive decline and may contribute to racial/ethnic disparities. Individuals eligible for KHANDLE were long-term members of KPNC who were age 65 or older on January 1, 2016, who previously participated in one or more Kaiser Permanente Multiphasic Health Checkups between 1964 and 1985, and who did not have a diagnosis of dementia in their electronic medical record. A total of 1,712 individuals were enrolled with efforts to recruit approximately equal proportions of Asian, Black, Latino, and White participants, including an overrepresentation of individuals with lower levels of educational attainment. The sample for this study consisted of 1,695 participants; 17 were excluded due to missing data. There were six who completed no cognitive tests, three did not identify as Asian, Black, Latino, or White, and eight were missing education. A randomly selected subgroup of 412 participants received comprehensive clinical evaluations with an adjudicated diagnosis and was used to evaluate associations of clinical diagnosis with cognitive tests. The study was approved by the KPNC and UC Davis Institutional Review Boards and all enrolled participants provided informed consent.

### Cognitive Assessment

#### Spanish and English Neuropsychological Assessment Scales

Three cognitive domains (verbal episodic memory, semantic memory, and executive function) were assessed by the SENAS, a battery of cognitive tests that has previously undergone extensive development for valid comparisons of cognitive change across diverse racial/ethnic groups and English and Spanish language administrations (Early et al., 2013; Mungas et al., 2000, 2004, 2005a, 2005b, 2010, 2011). Item response theory and confirmatory factor analysis methods were used to construct measures that are psychometrically matched across domains with respect to level of reliability across the ability continuum. Importantly, these measures do not have floor and ceiling effects and are normally distributed in the older adult population. The Episodic Memory score is derived from a multitrial word-list-learning test (Mungas et al., 2004). The Semantic Memory measure is a composite of highly correlated verbal (object-naming) and nonverbal (picture association) tasks. The Executive Function composite is constructed from component tasks of category fluency, phonemic (letter) fluency, and working memory (digit-span backward, list sorting; Crane et al., 2008). Administration procedures, measure development, and psychometric characteristics of

the SENAS battery are described in detail elsewhere (Crane et al., 2008; Mungas et al., 2004). SENAS measures used in this study were adjusted for differential item function (DIF) (Camilli and Shepard, 1994; Holland and Wainer, 1993) related to race/ethnicity, and this is described in [Supplementary Materials](#).

#### National Institutes of Health Toolbox Cognitive Health Battery

The NIHTB was conceived as an initiative to develop standardized measures of cognition, emotion, motor function, and sensation that could provide common research infrastructure to facilitate integration of results across studies (Gershon et al., 2013). The NIHTB-CHB measures multiple dimensions of cognition relevant to studies of cognitive function across the full range of normal cognition over an age range from 3 to 85 years (Weintraub et al., 2013). The development and validation of this battery is described in multiple publications (Gershon et al., 2013; Weintraub et al., 2013, 2014) and is presented in more detail in [Supplementary Materials](#).

Four measures from the NIHTB cognition domain were used in this study: the Flanker Inhibitory Control and Attention Test (Flanker), the Picture Sequence Memory Test (Picture Sequence Memory), the List Sorting Working Memory Test (List Sorting), and Picture Vocabulary. These four measures were administered using the iPad app for the NIHTB. Picture Sequence Memory measures episodic memory, Flanker and List Sorting measure components of executive function, and Picture Vocabulary measures language and semantic memory. The full battery of NIHTB-CHB cognitive tests was not administered due to participant burden and time constraints.

#### Language of cognitive assessment

Spanish and English language versions of both the SENAS and NIHTB-CHB were available. Language of test administration was determined by an algorithm that combined information regarding each participant's language preference in several specific contexts (e.g., conversing at home, listening to radio or television, conversing outside the home, preferred language for reading).

#### Clinical Diagnosis

##### Clinical evaluation design

The overall sample was divided into a 25% Random Selection group and a 75% Screen Selection group. Assignment to these groups was random within each racial/ethnic group so that each group had 25% randomly assigned to Random Selection. The Random Selection group was automatically invited to receive a clinical evaluation and results from this group were used for analyses related to clinical diagnosis ( $N = 412$ ). The SENAS and NIHTB-CHB measures used in this study were administered to all

participants in this group but were not considered in the process of establishing clinical diagnoses.

### Clinical evaluation components

The clinical evaluation consisted of two components: (a) clinical neuropsychological testing, and (b) clinical exam. Clinical neuropsychological testing was performed by a trained psychometrist, typically in the participant's home. The test battery included the neuropsychological test battery in version 3 of the Uniform Data set of the National Institute on Aging Alzheimer's Disease Centers program (Besser et al., 2018). Additional tests included the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) list learning test and CERAD drawing copy and delayed recall (Morris et al., 1989). Administration and scoring followed standard protocols. Test results were compared to two sets of norms to assess presence and pattern of cognitive impairment: one based on cognitively normal cases in the Random Selection group with no adjustment for demographic characteristics, and the second based on the same sample but with statistical adjustment for effects of race/ethnicity, gender, and education. More detail on adjustment methods is included in [Supplementary Materials](#). The clinical exam was administered by a physician specifically trained in clinical dementia assessment and was typically conducted in the participant's home. It included a medical history and history of cognitive complaints and problems, a physical and neurological exam, mental status testing using the Montreal Cognitive Assessment (MoCA; Nasreddine et al., 2005), and the Clinical Dementia Rating (CDR; Morris, 1993). Clinical evaluation and diagnosis are described in more detail in [Supplementary Materials](#).

### Diagnosis adjudication

Clinical diagnosis was adjudicated by a senior clinician (neurologist or neuropsychologist) with extensive experience with dementia assessment. This clinician reviewed the documented results of the clinical exam including the MoCA, the CDR, the medical history, and presentation of symptoms, and the examining clinician's impression of normal versus mild cognitive impairment (MCI) versus dementia. Neuropsychological test results and results of the neuropsychological adjudication were also available to the diagnostician. The diagnostician made sequential decisions: (a) overall classification (normal, MCI, dementia) based on clinical exam results and adjusted norm-referenced neuropsychological scores, blinded to demographic characteristics of the individual, and (b) final classification based on all available information including clinical exam results, participant demographics, raw neuropsychological test scores, and unadjusted and adjusted norm-referenced test scores. We used the final diagnosis based on all available information in this study. Reliability of clinical diagnosis is addressed in [Supplementary Materials](#).

## Data Analysis

### Measures and data processing

SENAS measures of Episodic Memory, Semantic Memory, and Executive Function and NIHTB-CHB measures Flanker, Picture Sequence Memory, List Sorting, and Picture Vocabulary were dependent variables. Independent variables included core demographic variables (age, gender, education), racial/ethnic group, and clinical diagnosis. Participants self-reported their race/ethnicity and were allowed to select multiple racial/ethnic categories. Those who identified with Latino/Hispanic ethnicity were classified as Latino in this study regardless of other categories that were selected, non-Latinos who identified as Black and other races were classified as Black, and Asians who also identified as White were classified as Asian. There were few cases of dementia, so clinical diagnosis was classified as impaired (MCI or dementia) versus normal. We applied a Blom transformation (Blom, 1958) to normalize cognitive variables and establish a common, standardized scale ( $M = 0$ ,  $SD = 1$ ). Years of education was centered at 12 years, and age in years was centered at 70 years. Gender and racial/ethnic group (in the combined sample analysis) were categorical variables coded using indicator variables. Racial/ethnic group was coded using three indicator variables: Black (1 = yes, 0 = no), Latino (1 = yes, 0 = no), and Asian (1 = yes, 0 = no); White was arbitrarily chosen as the reference group. Gender (male = 1, female = 0) was represented by single indicator variables. Clinical diagnosis was coded as an indicator variable (0 = normal, 1 = MCI or dementia).

### Demographic and diagnosis effect sizes

**Step 1:** We examined the simple bivariate associations of racial/ethnic group with cognitive measures in a combined sample of all four racial/ethnic groups. We estimated regression models for each cognitive outcome that included indicators for racial/ethnic groups as independent variables and calculated the variance explained ( $R^2$ ) by the racial/ethnic group indicators. **Step 2:** We then calculated incremental  $R^2$  explained by racial/ethnic group over that explained by demographic variables (age, education, gender, language of test administration), comparing a model with demographics plus racial/ethnic group indicators with a model including just demographics. We estimated variance explained by combined demographic variables and by racial/ethnic group. **Step 3:** Clinical diagnosis effects similarly were evaluated in the Random subsample using linear regression models that included the indicator variable for diagnosis (impaired vs normal), demographic variables, and indicator variables for racial/ethnic group as independent variables.

### Racial/ethnic differences in demographic and diagnosis effects

We utilized multiple group models to evaluate racial/ethnic group similarities and differences in effects of core demographic variables (age, education, gender)

and diagnosis on cognitive test scores. Language of test administration was not used in multiple group analyses because Spanish administration occurred only in the Latino group and in a small subsample ( $N = 48$ ). Multiple group analyses enable formal, hypothesis-driven tests of whether model parameters are invariant across groups. The general approach was to fit linear regression models in which regression coefficients for specific effects of interest were constrained to be equal across groups, and compare the fit of these constrained models with the fit of models in which these effects were allowed to differ across groups. Significantly better fit in the less constrained model indicates that model parameters differ across groups. Specifically, we examined group differences in core demographic variable effects on each cognitive outcome in the full sample as follows: **Step 1:** Separate analyses were performed for SENAS and NIHTB outcomes. **Step 2:** For each set of cognitive outcomes we started with a model in which estimated effects of age, education, and gender on each of the cognitive tests in the set were allowed to differ across racial/ethnic groups. **Step 3:** Subsequent models, one for each demographic variable, constrained the effects of that variable on individual cognitive variables to be the same across groups. The effects of that variable on different cognitive tests were allowed to differ, and the effects of other demographic variables were freely estimated as in Step 2. **Step 4:** The more constrained models from Step 3 were compared with the freely estimated models from Step 2 using the chi-square difference test and this provided an omnibus test of whether the specific demographic variable had different effects across groups for any of the cognitive variables in the set. **Step 5:** If the omnibus test showed a significant difference in model fit, then subsequent models constrained demographic effects to equality for one cognitive variable within the set at a time, and this model was compared with the Step 2 freely estimated model to determine if the demographic effect on that cognitive variable differed across groups. **Step 6:** If the significant group differences were found for individual cognitive tests, subsequent models constrained effects to be equal for pairs of groups to identify groups for which that demographic variable had differential effects on that cognitive variable. In summary, this process started with a family-wise, omnibus test that evaluated whether a given demographic variable had significantly different effects across groups for any cognitive variable and systematically performed more specific tests when the omnibus test was significant to identify the specific cognitive variables and groups that had differential effects. The process for clinical diagnosis was similar. The base model included demographic variables; demographic effects for variables that were found to have differential effects on cognitive scores in previous analyses were freely estimated across groups and those that did not have differential effects were constrained to equality. Significance

of group differences in incremental diagnosis effects was estimated by comparing a model in which diagnosis effects were freely estimated with a model in which they were constrained to equality.

Multiple group analyses were performed using Mplus version 8.2 (Muthén & Muthén, 1998) and descriptive statistics and demographic and diagnosis effect size analyses were performed using R version 3.6.1.

## Results

### Sample Characteristics

Sample characteristics are presented in Table 1, stratified by racial/ethnic group. About 59% of the sample were females. Gender representation differed across groups ( $\chi^2[3] = 19.311, p = .001$ )—Blacks had the highest representation of women (67.4%) and Asians the lowest (53%). Average age was about 76 years and while differences across groups were significant ( $F[3,1691] = 3.608, p = .013$ ), average differences were 1.4 years at most. Average education was 14.6 years and differed across groups ( $F[3,1691] = 50.251, p = .001$ ). Spanish administration of cognitive tests occurred in 48 individuals, all Latinos. All others were tested in English. About 1% of the Random Selection group were diagnosed with dementia, 20% had MCI, and 79% were cognitively normal. Diagnosis differed by race/ethnicity ( $\chi^2[6] = 18.267, p = .006$ ) with Whites more likely to be normal. There were significant racial/ethnic group differences in average scores for the SENAS and NIHTB cognitive measures ( $ps < .001$ ).

### Demographic Effects in Combined Sample

Table 2 shows strength of association ( $R^2$ ) of racial/ethnic group with individual cognitive measures. Racial/ethnic group explained from 2% to 20% of the variance in simple bivariate associations. Incremental effects of racial/ethnic group above and beyond the effects of core demographic variables (age, gender, education, language of test administration) were of similar magnitude as the simple effects. The SENAS and NIHTB measures of episodic memory were least associated with racial/ethnic group and measures of semantic memory/vocabulary were most affected. Racial/ethnic group and demographic effects combined explained 15%–40% of the variance.

Figure 1 shows the associations of core demographic variables on cognitive test scores in the combined sample. Older age was negatively related to all cognitive measures, and with the exception of NIHTB Picture Vocabulary, magnitude of effects was similar. Education was related to all measures, with strongest associations for Picture Vocabulary and SENAS Semantic Memory. Males had lower scores on most cognitive tests, especially measures of episodic memory, but females had lower scores on SENAS Semantic Memory.

**Table 1.** Sample Characteristics

	Asian	Black	Latino	White	Total
Gender—female <sup>a</sup>	219 (53.0%)	296 (67.4%)	202 (58.7%)	288 (57.7%)	1,005 (59.3%)
Age (years)—mean ( <i>SD</i> )	75.3 ( $\pm 7.0$ )	75.1 ( $\pm 7.1$ )	75.6 ( $\pm 6.6$ )	76.5 ( $\pm 7.5$ )	75.6 ( $\pm 7.1$ )
Education (years)—mean ( <i>SD</i> )	15.6 ( $\pm 2.6$ )	14.2 ( $\pm 2.8$ )	13.1 ( $\pm 4.0$ )	15.2 ( $\pm 2.9$ )	14.6 ( $\pm 3.2$ )
Language of Test Administration—Spanish <sup>a</sup>	0 (0.0%)	0 (0.0%)	48 (14.0%)	0 (0.0%)	1,647 (97.2%)
Clinical Diagnosis—normal <sup>a</sup>	80 (78.4%)	63 (71.6%)	88 (75.9%)	93 (87.7%)	324 (78.6%)
Clinical Diagnosis—MCI <sup>a</sup>	22 (21.6%)	21 (23.9%)	28 (24.1%)	12 (11.3%)	83 (20.1%)
Clinical Diagnosis—dementia <sup>a</sup>	0 (0.0%)	4 (4.5%)	0 (0.0%)	1 (0.9%)	5 (1.2%)
Episodic Memory (SENAS)—mean ( <i>SD</i> )	0.2 ( $\pm 1.0$ )	-0.1 ( $\pm 0.9$ )	-0.2 ( $\pm 1.0$ )	0.1 ( $\pm 1.0$ )	0.0 ( $\pm 1.0$ )
Semantic Memory (SENAS)—mean ( <i>SD</i> )	-0.2 ( $\pm 1.0$ )	-0.5 ( $\pm 0.8$ )	0.0 ( $\pm 0.9$ )	0.6 ( $\pm 0.9$ )	0.0 ( $\pm 1.0$ )
Executive Function (SENAS)—mean ( <i>SD</i> )	-0.2 ( $\pm 0.9$ )	-0.3 ( $\pm 0.9$ )	-0.2 ( $\pm 0.9$ )	0.5 ( $\pm 1.1$ )	0.0 ( $\pm 1.0$ )
Flanker (NIHTB)—mean ( <i>SD</i> )	0.2 ( $\pm 1.0$ )	-0.4 ( $\pm 1.0$ )	0.0 ( $\pm 1.0$ )	0.2 ( $\pm 0.9$ )	0.0 ( $\pm 1.0$ )
Picture Sequence Memory (NIHTB)—mean ( <i>SD</i> )	0.0 ( $\pm 1.0$ )	-0.3 ( $\pm 0.9$ )	0.1 ( $\pm 0.9$ )	0.2 ( $\pm 1.0$ )	0.0 ( $\pm 1.0$ )
List Sorting (NIHTB)—mean ( <i>SD</i> )	0.0 ( $\pm 0.9$ )	-0.3 ( $\pm 1.0$ )	0.0 ( $\pm 0.9$ )	0.2 ( $\pm 1.0$ )	0.0 ( $\pm 1.0$ )
Picture Vocabulary (NIHTB)—mean ( <i>SD</i> )	-0.1 ( $\pm 1.0$ )	-0.4 ( $\pm 0.9$ )	-0.2 ( $\pm 0.9$ )	0.6 ( $\pm 0.9$ )	0.0 ( $\pm 1.0$ )

Notes: MCI = mild cognitive impairment; NIHTB = NIH Toolbox; *SD* = standard deviation; SENAS = Spanish and English Neuropsychological Assessment Scales. Results are from the Kaiser Healthy Aging and Diverse Life Experiences study ( $N = 1,695$ ).

<sup>a</sup>Percents in parentheses are within racial/ethnic group percentages.

**Table 2.** Strength of Association ( $R^2$ ) of Racial/Ethnic Group and Core Demographic Variables With Cognitive Measures in Full Sample ( $N = 1,695$ )

Cognitive measure	Racial/ethnic group (simple) <sup>a</sup>	Demographic	Group + demographic	Racial/ethnic group (incremental) <sup>b</sup>
Episodic Memory (SENAS)	0.02	0.25	0.26	0.02
Semantic Memory (SENAS)	0.20	0.21	0.40	0.19
Executive Function (SENAS)	0.11	0.21	0.35	0.14
Flanker (NIHTB)	0.06	0.14	0.20	0.06
Picture Sequence Memory (NIHTB)	0.03	0.12	0.15	0.03
List Sorting (NIHTB)	0.04	0.14	0.21	0.07
Picture Vocabulary (NIHTB)	0.14	0.24	0.35	0.11

Notes: NIHTB = NIH Toolbox; SENAS = Spanish and English Neuropsychological Assessment Scales.

<sup>a</sup>“Simple” racial/ethnic group effects refer to bivariate associations with cognitive variables.

<sup>b</sup>“Incremental” racial/ethnic group is the difference between models with core demographic variables and core demographic variables plus racial/ethnic group.

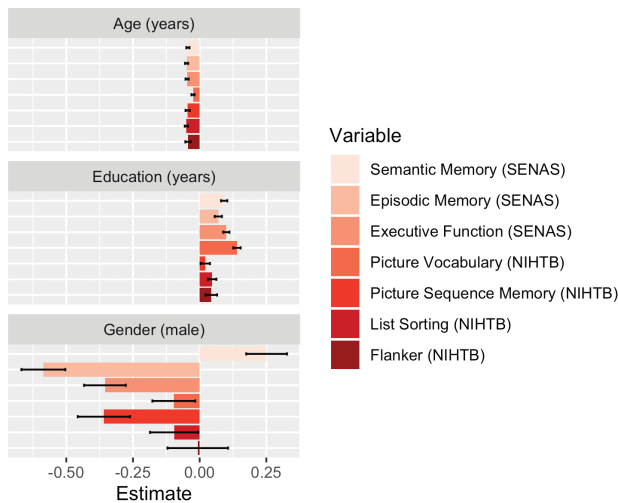
### Diagnosis Effects in Combined Sample

Associations of cognitive test results with a clinical diagnosis of cognitive impairment are presented in Figure 2. These results show incremental clinical diagnosis effects independent of racial/ethnic group, age, education, and gender. Cognitive impairment was associated with substantially lower scores for all seven cognitive measures. The cognitive impairment effect was largest for SENAS Executive Function; those diagnosed with cognitive impairment scored 0.868 *SD* lower on average than those diagnosed as normal. SENAS Episodic Memory and NIHTB List Sorting and Flanker had cognitive impairment effect sizes exceeding  $-0.5$  *SD*, and the smallest effect size ( $-0.415$  *SD*) was for NIHTB Picture Vocabulary.

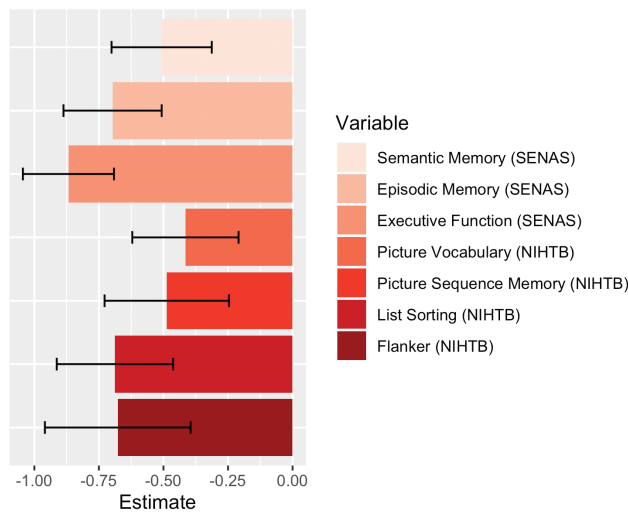
### Demographic and Diagnosis Effect Differences by Racial/Ethnic Group

Effects of demographic variables on SENAS scores did not significantly differ across racial/ethnic groups (age:

$\chi^2[9] = 10.924, p = .281$ ; gender:  $\chi^2[9] = 8.715, p = .464$ ; education:  $\chi^2[9] = 15.461, p = .079$ ). For the NIHTB, gender ( $\chi^2[12] = 31.506, p = .002$ ) and education ( $\chi^2[12] = 32.239, p = .001$ ) effects significantly differed across racial/ethnic groups, but age effects did not differ ( $\chi^2[12] = 10.193, p = .599$ ). The gender effect differed across groups for Picture Vocabulary ( $\chi^2[3] = 13.985, p = .003$ ) and List Sorting ( $\chi^2[3] = 9.667, p = .022$ ) but not for Flanker ( $p = .053$ ) or Picture Sequence Memory ( $p = .525$ ). For Picture Vocabulary, Black and Latino males had significantly lower scores than females. For List Sorting, Latino males scored significantly higher than females while Asian and White males had lower average scores in comparison with females. The education effect differed across groups for Flanker ( $\chi^2[3] = 8.99, p = .029$ ) and List Sorting ( $\chi^2[3] = 25.118, p = .001$ ) but not for Picture Sequence Memory ( $p = .657$ ) or Picture Vocabulary ( $p = .12$ ). Education effects on Flanker and List Sorting were not significant for Whites but were significant for the three other groups.



**Figure 1.** Associations of core demographic variables with SENAS and NIHTB cognition measures. Colored bars (with black 95% confidence interval bars) represent effect estimates for cognitive test scores regressed on age, education, and gender in the full sample. Effects for age and education show the average impact in standard deviation units of a 1 year increment in age or education. The effects for gender show how the average scores for males differ from those for females. Results are from models that included all core demographic variables as independent variables. NIHTB = NIHToolbox; SENAS = Spanish and English Neuropsychological Assessment Scales.



**Figure 2.** Associations of clinical diagnosis with cognitive test scores in Random Selection subgroup. Colored bars (with black 95% confidence interval bars) show the magnitude of average differences between cognitively impaired individuals (MCI or dementia) in comparison with cognitively normal individuals. Results are from models that included all core demographic variables and indicators for racial/ethnic group as independent variables. MCI = mild cognitive impairment; NIHTB = NIH Toolbox; SENAS = Spanish and English Neuropsychological Assessment Scales.

Table 3 shows the effects of clinical diagnosis by racial/ethnic group. The pattern of diagnosis differences did not significantly differ across groups for SENAS ( $\chi^2[9] = 13.546, p = .139$ ) or for NIHTB-CHB measures ( $\chi^2[12] = 11.687, p = .471$ ). While there were some differences in freely

estimated diagnosis effects across racial/ethnic groups for some cognitive measures, the overall pattern of differences was not statistically significant in the omnibus tests and there were striking similarities across groups. Four of the seven cognitive measures (SENAS Episodic Memory and Executive Function, NIHTB-CHB Flanker and List Sorting) showed significant associations with diagnosis in all four groups and SENAS Semantic Memory and NIHTB Picture Vocabulary significantly differed by Diagnosis in three groups. NIHTB-CHB Picture Sequence Memory was associated with diagnosis in Whites and Latinos.

### Discussion

Results showed robust effects of age and education on SENAS and NIHTB-CHB measures of different cognitive domains. Racial/ethnic group differences varied substantially across cognitive measures, and were smallest for tests of episodic and working memory and largest for vocabulary and semantic memory. Females generally had higher average scores, with a notable exception that males had higher SENAS semantic memory scores. Effects of demographic variables on SENAS scores did not significantly differ across racial/ethnic groups. For the NIHTB-CHB, gender and education effects differed across groups but age effects did not differ. Clinical diagnosis (normal vs impaired) had strong effects in the Random Selection subgroup on all variables independent of racial/ethnic group and core demographic variables. Diagnosis effects were strongest for measures of episodic memory, executive function, and working memory, and were weakest for vocabulary and semantic memory. Overall, clinical diagnosis effects on test scores did not significantly differ across racial/ethnic groups, and with a few exceptions, clinical diagnosis had clinically and statistically significant effects on test scores within specific groups.

The pattern of demographic and racial/ethnic group effects on cognitive test score is generally consistent with previous literature (Brewster et al., 2014; Cagney & Lauderdale, 2002; Castora-Binkley et al., 2015; Early et al., 2013; Gross et al. 2015; Masel & Peek, 2009; Mungas et al., 2009; Weuve et al., 2018; Wilson et al., 2016). Education and racial/ethnic group had strongest effects on measures of vocabulary and semantic memory that reflect acquired knowledge over the life span, and weakest effects on episodic memory that reflects ability to learn and retain new information. Clinical diagnosis was most strongly related to episodic memory and executive function and least related to semantic memory and vocabulary. This is consistent with previous studies that showed episodic memory to be most strongly related to diagnosis and brain variables and executive function to have stronger associations than semantic memory with brain integrity measures (Mungas et al., 2009, 2010). These results indicate that episodic memory and executive function are better indicators of clinical status and brain health than

**Table 3.** Effects of Clinical Diagnosis by Racial/Ethnic Group

Dependent variable	Asian	Black	Latino	White
Episodic Memory	-0.408 (0.177)*	-0.741 (0.179)***	-0.819 (0.233)***	-0.736 (0.156)***
Semantic Memory	-0.331 (0.161)*	-0.570 (0.217)**	-0.694 (0.198)***	-0.271 (0.141)
Executive Function	-0.945 (0.152)***	-0.562 (0.187)**	-1.133 (0.216)***	-0.804 (0.133)***
Flanker	-0.600 (0.258)*	-0.882 (0.247)***	-0.782 (0.254)**	-0.729 (0.282)**
Picture Sequence Memory	-0.139 (0.224)	-0.379 (0.235)	-0.619 (0.278)*	-0.731 (0.184)***
List Sorting	-0.709 (0.207)***	-0.426 (0.204)*	-0.942 (0.306)**	-0.467 (0.186)*
Picture Vocabulary	-0.407 (0.168)*	-0.081 (0.253)	-0.449 (0.213)*	-0.562 (0.162)***

Notes: Results show the effects of clinical diagnosis (normal vs impaired) on cognitive variables for each racial/ethnic group. Tabled values show how the average scores for cognitively impaired individuals within a group (mild cognitive impairment or dementia) differ from those for cognitively normal individuals within that group. Results are from multiple group models that freely estimated diagnosis effects on cognitive variables across groups.

\* $p < .05$ . \*\* $p < .01$ . \*\*\* $p < .001$ .

vocabulary/semantic memory. A unique contribution is that this is the first study, to our knowledge, to examine these effects in four major racial/ethnic groups using comprehensive measures of cognition.

Results showed that the effects of demographic variables and clinical diagnosis did not differ across racial/ethnic groups, with minor exceptions, notably that education effects on two NIHTB measures (Flanker and List Sorting) were weaker in Whites. These differences are consistent with [Jean et al. \(2019\)](#), who found greater education effects in Blacks compared to Whites on nonepisodic memory domains. However, in this study, education effects on all three SENAS measures and two of the four NIHTB measures did not differ by racial/ethnic group. This is important evidence that cognitive test results have the same associations with important external variables across these very diverse groups, and this is relevant for using cognitive tests to characterize and study cognitive health and cognitive aging in the increasingly diverse older adult population. The robust effects of clinical diagnosis independent of racial/ethnic group and core demographics provide strong evidence of the validity of cognitive tests as measures of cognitive health.

There were substantial similarities of results for SENAS and NIHTB measures, but there also were some differences, notably that some NIHTB measures were differentially affected across racial/ethnic groups by education and gender. One relevant difference between these two sets of cognitive measures was that SENAS development yielded scores that were adjusted for differential item function related to race/ethnicity. DIF adjustment should decrease measurement bias, and it is possible that the SENAS associations with demographic variables better reflect true associations within these groups. However, SENAS and NIHTB-CHB tests differ in many other respects that could also explain the differences in results. SENAS development has occurred over nearly three decades and many studies have addressed validity for identifying and monitoring cognitive impairment in older adults who are diverse in race/ethnicity, educational attainment, linguistic and cultural background, and presence and severity of AD

and related diseases of aging ([Early et al. 2013](#); [Mungas et al., 2000, 2004, 2005a, 2005b, 2010](#)). The NIHTB is a newer effort that was designed to measure normal range function ([Weintraub et al., 2013](#)); validation in clinical groups was not part of the original development plan. This study represents an important step toward establishing a rigorous empirical base for understanding how cognition is affected by clinical disease and demographic diversity across comprehensive cognitive measures from different development pipelines.

A major strength of this study was the availability of four racial/ethnic groups and the relatively large sample size for each of these groups. The sample sizes, ranging from 344 to 499 participants per group, provide good statistical power for detecting differential effects of age, education, and gender across groups. Another major strength is the comprehensive and multimethod approach to measuring cognition. The clinical adjudication protocol was a rigorous, standardized, and quantitative process involving expert neurologists and neuropsychologists. The clinical adjudication excluded consideration of SENAS and NIHTB-CHB, thus avoiding a major source of methodological bias. Specifically, diagnosis was based on an entirely different cognitive test battery that was administered on a different date. However, due to the smaller sample size of the Random Selection subgroup, statistical power for detecting group difference in diagnosis effects was lower than that for detecting group differences in demographic effects in the full sample. Nevertheless, the results still showed significant effects of diagnosis in all groups for four of the seven cognitive measures.

The main limitation is that this study only examined cross-sectional baseline data, and therefore cognitive health could be directly measured as change across sequential measurements. In cross-sectional assessments, cognitive health has to be inferred by comparing observed test performance with normative data of healthy individuals with similar background characteristics. Given that there were robust effects of clinical diagnosis independent of racial/ethnic group and demographic variables in the full sample, and that diagnosis



effects were present in all four groups independent of demographics, this study makes a compelling case for using cognitive test scores as indicators for brain health in other clinical and research applications involving diverse populations. Sampling bias is a second potential limitation. This sample included individuals who have been enrolled in the KPNC for decades and so have had unusual access to health care. In addition, educational achievement, particularly for males, was higher than the general population. However, there was a broad range of education attainment in the sample and this should facilitate sensitive estimation of education associations with cognitive test scores. A third limitation is that this study only examined the effects of age, gender, and education and did not include other external variables such as early life sociocultural experiences and literacy level, which previous literature has shown to be significantly associated with cognitive test scores in late life (Brewster et al., 2014; Manly et al., 2002; Sisco et al., 2015). Looking ahead, data from subsequent KHANDLE assessment waves, once available, will provide the ability to measure change in cognitive health from repeated assessments and disentangle which variables are and are not associated with change in brain health.

The results of this study expand literature on demographic and clinical effects on cognitive test scores in four major racial/ethnic groups. The overall pattern of results shows similar effects of demographic variables and diagnosis across racial/ethnic groups. Results provide important information for clinical and research applications of cognitive tests to characterize, monitor, and study cognitive health in older adults.

## Supplementary Material

Supplementary data are available at *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences* online.

## Funding

This work was supported by a grant from the National Institute on Aging (NIA) (RF1 AG052132, R. A. Whitmer, D. Mungas, C. DeCarli, M. Glymour PIs).

## Conflict of Interest

None declared.

## Acknowledgments

The raw data that support the findings of this study are available from the senior author upon request, subject to establishing a data use agreement. This study was not preregistered.

## References

- Besser, L., Kukull, W., Knopman, D. S., Chui, H., Galasko, D., Weintraub, S., Jicha, G., Carlsson, C., Burns, J., Quinn, J., Sweet, R. A., Rascovsky, K., Teylan, M., Beekly, D., Thomas, G., Bollenbeck, M., Monsell, S., Mock, C., Zhou, X. H., . . . Clinical Core leaders of the National Institute on Aging. (2018). Version 3 of the National Alzheimer's Coordinating Center's Uniform Data Set. *Alzheimer Disease and Associated Disorders*, 32(4), 351–358. doi:10.1097/WAD.0000000000000279
- Blom, G. 1958. *Statistical estimates and transformed beta-variables*. Wiley.
- Brewster, P. W., Melrose, R. J., Marquine, M. J., Johnson, J. K., Napoles, A., MacKay-Brandt, A., Farias, S., Reed, B., & Mungas, D. (2014). Life experience and demographic influences on cognitive function in older adults. *Neuropsychology*, 28(6), 846–858. doi:10.1037/neu0000098
- Byrd, D. R., Gee, G. C., & Tarraf, W. (2018). Black–White mental status trajectories: What ages do differences emerge? *SSM—Population Health*, 6(December), 169–177. doi:10.1016/j.ssmph.2018.09.008
- Cagney, K. A., & Lauderdale, D. S. (2002). Education, wealth, and cognitive function in later life. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 57(2), P163–P172. doi:10.1093/geronb/57.2.p163
- Camilli, G., & Shepard, L. A. 1994. *Methods for identifying biased test items* (Vol. 4). Sage Publications, Inc.
- Castora-Binkley, M., Peronto, C. L., Edwards, J. D., & Small, B. J. (2015). A longitudinal analysis of the influence of race on cognitive performance. *The Journals of Gerontology. Series B, Psychological Sciences and Social Sciences*, 70(4), 512–518. doi:10.1093/geronb/gbt112
- Cotter V. T. (2007). The burden of dementia. *The American Journal of Managed Care*, 13(Suppl 8), S193–S197.
- Crane, P. K., Narasimhalu, K., Gibbons, L. E., Pedraza, O., Mehta, K. M., Tang, Y., Manly, J. J., Reed, B. R., & Mungas, D. M. (2008). Composite scores for executive function items: Demographic heterogeneity and relationships with quantitative magnetic resonance imaging. *Journal of the International Neuropsychological Society*, 14(5), 746–759. doi:10.1017/S1355617708081162
- Díaz-Venegas, C., Downer, B., Langa, K. M., & Wong, R. (2016). Racial and ethnic differences in cognitive function among older adults in the USA. *The International Journal of Geriatric Psychiatry*, 31(9), 1004–1012. doi:10.1002/gps.4410
- Dowling, N. M., Farias, S. T., Reed, B. R., Sonnen, J. A., Strauss, M. E., Schneider, J. A., Bennett, D. A., & Mungas, D. (2011). Neuropathological associates of multiple cognitive functions in two community-based cohorts of older adults. *Journal of the International Neuropsychological Society*, 17(4), 602–614. doi:10.1017/S1355617710001426
- Early, D. R., Widaman, K. F., Harvey, D., Beckett, L., Park, L. Q., Farias, S. T., Reed, B. R., Decarli, C., & Mungas, D. (2013). Demographic predictors of cognitive change in ethnically diverse older persons. *Psychology and Aging*, 28(3), 633–645. doi:10.1037/a0031645
- Fjell, A. M., McEvoy, L., Holland, D., Dale, A. M., & Walhovd, K. B. (2014). What is normal in normal aging? Effects of aging, amyloid and Alzheimer's disease on the cerebral cortex and the hippocampus. *Progress in Neurobiology*, 117, 20–40. doi:10.1016/j.pneurobio.2014.02.004

- Fletcher, E., Gavett, B., Harvey, D., Farias, S. T., Olichney, J., Beckett, L., DeCarli, C., & Mungas, D. (2018). Brain volume change and cognitive trajectories in aging. *Neuropsychology*, 32(4), 436–449. doi:10.1037/neu0000447
- Gershon, R. C., Wagster, M. V., Hendrie, H. C., Fox, N. A., Cook, K. F., & Nowinski, C. J. (2013). NIH Toolbox for assessment of neurological and behavioral function. *Neurology*, 80(11 Suppl. 3), S2–S6. doi:10.1212/WNL.0b013e3182872e5f
- Glymour, M. M., & Manly, J. J. (2008). Lifecourse social conditions and racial and ethnic patterns of cognitive aging. *Neuropsychology Review*, 18(3), 223–254. doi:10.1007/s11065-008-9064-z
- Gross, A. L., Mungas, D. M., Crane, P. K., Gibbons, L. E., MacKay-Brandt, A., Manly, J. J., Mukherjee, S., Romero, H., Sachs, B., Thomas, M., Potter, G. G., & Jones, R. N. (2015). Effects of education and race on cognitive decline: An integrative study of generalizability versus study-specific results. *Psychology and Aging*, 30(4), 863–880. doi:10.1037/pag0000032
- Holland, P. W., & Wainer, H. (1993). *Differential item functioning: Theory and practice*. Lawrence Erlbaum.
- Jean, K. R., Lindbergh, C. A., Mewborn, C. M., Robinson, T. L., Gogniat, M. A., & Stephen Miller, L. (2019). Education differentially buffers cognitive performance in Black and White older adults. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 74(8), 1366–1375. doi:10.1093/geronb/gby116
- Karlamangla, A. S., Miller-Martinez, D., Aneshensel, C. S., Seeman, T. E., Wight, R. G., & Chodosh, J. (2009). Trajectories of cognitive function in late life in the United States: demographic and socioeconomic predictors. *American Journal of Epidemiology*, 170(3), 331–342. doi:10.1093/aje/kwp154
- Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society: JINS*, 8(3), 341–348. doi:10.1017/s1355617702813157
- Masel, M. C., & Peek, M. K. (2009). Ethnic differences in cognitive function over time. *Annals of Epidemiology*, 19(11), 778–783. doi:10.1016/j.annepidem.2009.06.008
- Matthews, K. A., Xu, W., Gaglioti, A. H., Holt, J. B., Croft, J. B., Mack, D., & McGuire, L. C. (2019). Racial and ethnic estimates of Alzheimer's disease and related dementias in the United States (2015–2060) in adults aged ≥65 years. *Alzheimer's and Dementia*, 15(1), 17–24. doi:10.1016/j.jalz.2018.06.3063
- Melrose, R. J., Brewster, P., Marquine, M. J., MacKay-Brandt, A., Reed, B., Farias, S. T., & Mungas, D. (2015). Early life development in a multiethnic sample and the relation to late life cognition. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 70(4), 519–531. doi:10.1093/geronb/gbt126
- Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, 43(11), 2412–2414. doi:10.1212/wnl.43.11.2412-a
- Morris, J. C., Heyman, A., Mohs, R. C., Hughes, J. P., van Belle, G., Fillenbaum, G., Mellits, E. D., & Clark, C. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, 39(9), 1159–1165. doi:10.1212/wnl.39.9.1159
- Mungas, D., Beckett, L., Harvey, D., Farias, S. T., Reed, B., Carmichael, O., Olichney, J., Miller, J., & DeCarli, C. (2010). Heterogeneity of cognitive trajectories in diverse older persons. *Psychology and Aging*, 25(3), 606–619. doi:10.1037/a0019502
- Mungas, D., Reed, B. R., Crane, P. K., Haan, M. N., & González, H. (2004). Spanish and English Neuropsychological Assessment Scales (SENAS): Further development and psychometric characteristics. *Psychological Assessment*, 16(4), 347–359. doi:10.1037/1040-3590.16.4.347
- Mungas, D., Reed, B. R., Farias, S., & DeCarli, C. (2005a). Criterion-referenced validity of a neuropsychological test battery: Equivalent performance in elderly Hispanics and non-Hispanic Whites. *Journal of the International Neuropsychological Society*, 11(5), 620–630. doi:10.1017/S1355617705050745
- 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. doi:10.1037/a0013421
- Mungas, D., Reed, B. R., Haan, M. N., & González, H. (2005b). Spanish and English neuropsychological assessment scales: Relationship to demographics, language, cognition, and independent function. *Neuropsychology*, 19(4), 466–475. doi:10.1037/0894-4105.19.4.466
- Mungas, D., Reed, B. R., Marshall, S. C., & González, H. M. (2000). Development of psychometrically matched English and Spanish language neuropsychological tests for older persons. *Neuropsychology*, 14(2), 209–223. doi:10.1037//0894-4105.14.2.209
- Mungas, D., Widaman, K. F., Reed, B. R., & Farias, S. T. (2011). Measurement invariance of neuropsychological tests in diverse older persons. *Neuropsychology*, 25(2), 260–269. doi:10.1037/a0021090
- Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide* (8th ed.). Author.
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53(4), 695–699. doi:10.1111/j.1532-5415.2005.53221.x
- 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. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 70(4), 557–567. doi:10.1093/geronb/gbt133
- Treas, J., & Carreon, D. (2010). Diversity and our common future: Race, ethnicity, and the older American. *Generations*, 34(3), 38–44.
- Walter, S., Dufouil, C., Gross, A. L., Jones, R. N., Mungas, D., Filshtein, T. J., Manly, J. J., Arpawong, T. E., & Glymour, M. M. (2019). Neuropsychological test performance and MRI markers of dementia risk: Reducing education bias. *Alzheimer Disease and Associated Disorders*, 33(3), 179–185. doi:10.1097/WAD.0000000000000321
- Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Bauer, P. J., Carlozzi, N. E., Slotkin, J., Blitz, D.,

- Wallner-Allen, K., Fox, N. A., Beaumont, J. L., Mungas, D., Nowinski, C. J., Richler, J., Deocampo, J. A., Anderson, J. E., Manly, J. J., Borosh, B., ... Gershon, R. C. (2013). Cognition assessment using the NIH Toolbox. *Neurology*, *80*(11 Suppl. 3), S54–S64. doi:[10.1212/WNL.0b013e3182872ded](https://doi.org/10.1212/WNL.0b013e3182872ded)
- Weintraub, S., Dikmen, S. S., Heaton, R. K., Tulsky, D. S., Zelazo, P. D., Slotkin, J., Carlozzi, N. E., Bauer, P. J., Wallner-Allen, K., Fox, N., Havlik, R., Beaumont, J. L., Mungas, D., Manly, J. J., Moy, C., Conway, K., Edwards, E., Nowinski, C. J., & Gershon, R. (2014). The cognition battery of the NIH toolbox for assessment of neurological and behavioral function: Validation in an adult sample. *Journal of the International Neuropsychological Society*, *20*(6), 567–578. doi:[10.1017/S1355617714000320](https://doi.org/10.1017/S1355617714000320)
- Weuve, J., Barnes, L. L., Mendes de Leon, C. F., Rajan, K. B., Beck, T., Aggarwal, N. T., Hebert, L. E., Bennett, D. A., Wilson, R. S., & Evans, D. A. (2018). Cognitive aging in Black and White Americans: Cognition, cognitive decline, and incidence of Alzheimer disease dementia. *Epidemiology (Cambridge, Mass.)*, *29*(1), 151–159. doi:[10.1097/EDE.0000000000000747](https://doi.org/10.1097/EDE.0000000000000747)
- Wilson, R. S., Rajan, K. B., Barnes, L. L., Weuve, J., & Evans, D. A. (2016). Factors related to racial differences in late-life level of cognitive function. *Neuropsychology*, *30*(5), 517–524. doi:[10.1037/neu0000290](https://doi.org/10.1037/neu0000290)