The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences, 2025, **80**(S1), gbae036
https://doi.org/10.1093/geronb/gbae036
Advance access publication 5 April 2024
Supplement Article: Studying Cognition in the National Social Life, Health, and Aging Project: Methodological and Analytic Considerations, Strategies and Findings From Round 4

GSA GERONTOLOGICAL SOCIETY OF AMERICA®
OXFORD

Addressing Cognitive Assessment Disparities Among Hispanic Adults: Adapting the MoCA-SA for Improved Accuracy and Accessibility Among Spanish Speakers

Lissette M. Piedra, PhD, MSW,[1,*] James Iveniuk, PhD,[2] Melissa J.K. Howe, PhD,[2] Kelly Pudelek, MA,[2] and David X. Marquez, PhD[3]

[1]School of Social Work, University of Illinois at Urbana-Champaign, Urbana, Illinois, USA.
[2]The Bridge at NORC at the University of Chicago, Chicago, Illinois, USA.
[3]Department of Kinesiology and Nutrition, University of Illinois Chicago, Chicago, Illinois, USA.

*Address correspondence to: Lissette M. Piedra, PhD, MSW. E-mail: lmpiedra@illinois.edu

Decision Editor: Hui Liu, PhD (Social Sciences Section)

Abstract

Objectives: Hispanic adults display a higher likelihood of early-stage cognitive decline than their White counterparts yet receive fewer clinical diagnoses. This troubling trend highlights the significance of longitudinal surveys like the National Social Life, Health, and Aging Project (NSHAP) in monitoring cognitive changes in aging Hispanics. Using NSHAP's Rounds 2 and 3, we observed notable cognitive score differences between English and Spanish speakers, as assessed by the survey-adapted version of the Montreal Cognitive Assessment (MoCA-SA). Our study aims to discern if statistical adjustments can reduce measurement variance in global cognition scores between these language groups.

Methods: We applied modification indexes, 2-parameter item response theory models, and split-sample testing to pinpoint items that exhibit resilience to language-related bias among our Hispanic sample. From this analysis, an abbreviated version of the MoCA-SA, termed MoCA-SAA, was introduced. Subsequently, we juxtaposed the performance and predictive validity of both MoCA versions against four consequential outcomes indicative of cognitive decline.

Results: Our refined methodologies enabled the identification of consistent items across both language cohorts. The MoCA-SAA demonstrated performance and predictive validity in line with the original MoCA-SA concerning outcomes linked to cognitive deterioration.

Discussion: The translated measures ensure the inclusion of Hispanic Spanish speakers in NSHAP, who might otherwise be overlooked. The statistical adjustment outlined in this study offers a means to mitigate potential measurement disparities when assessing overall cognition. Despite these advancements, we acknowledge persistent issues related to the translation of the MoCA-SA into Spanish that warrant further attention.

Keywords: Cognitive health, Hispanic older adults, MoCA, Spanish-speaking older adults

The United States has seen a significant surge in its Spanish-speaking population, with an increase of 30.6 million speakers from 1980 to 2019, encompassing those fluent in Spanish Creole and Ladino (Dietrich & Hernandez, 2022). Within this growing Spanish-speaking demographic, a substantial majority (94%) identify as Hispanic, with a predominant racial affiliation of White (65%), followed by a significant presence in the "some other race" category (32.3%). These statistics highlight a notable and evolving demographic trend.

As the Hispanic/Latino/a/x/e population, henceforth referred to as Hispanic, continues to age, there is a significant projection of growth among Hispanic individuals aged 65 and older. According to data from the Administration for Community Living (2020), this population is expected to increase from nearly 5 million in 2020 to a remarkable 19.9 million by 2060. This demographic shift also applies to the subset of Hispanic adults facing Alzheimer's disease and related dementias (ADRD). Estimates of ADRD prevalence within this group vary, ranging from 5% to 12.9%, depending on factors like the choice of cognitive assessments, the age composition of the sample, and diverse cultural heritages (Angel & Rote, 2021; González et al., 2019; Haan et al., 2003; Matthews et al., 2019; Mehta & Yeo, 2017). Notably, one projection anticipates an astonishing 832% increase, forecasting the number of Hispanic individuals living with ADRD to rise from 379,000 in 2012 to 3.5 million by 2060 (Wu et al., 2016). While many Hispanic adults are bilingual (Krogstad & Gonzalez-Barrera, 2015), 21% of those ages 65 and older are monolingual in Spanish and do not speak English (Krogstad et al., 2015).

The limited availability of services in Spanish in the United States exacerbates language barriers, hindering access to necessary services. A substantial body of literature demonstrates the problems and negative consequences stemming from

Received: May 15 2023; Editorial Decision Date: December 15 2023.

language mismatches between patients and the healthcare system (Bauer et al., 2010; Ginde et al., 2009; Piedra et al., 2011; Pot et al., 2020; Schenker et al., 2007; Schyve, 2007; Showstack, 2019; Terui & Hsieh, 2022; Zhao et al., 2021). Language barriers can impede access to ancillary services, hinder compliance with medical recommendations, and create an overreliance on bilingual caregivers that deters their use of respite services.

Language barriers can also complicate measures of cognitive health and the detection of cognitive decline as Hispanics age. For example, ADRD may accelerate language reversion, that is, a shift back to their native language that can occur with age (Goth & Strøm, 2018; Massett et al., 2021; McMurtray et al., 2009). Hispanic adults are more likely to develop mild cognitive impairment and dementia at earlier ages than their White peers (Angel & Rote, 2021; Fitten et al., 2014; Mehta & Yeo, 2017), and they are among the least likely to be clinically diagnosed with early-stage dementia or mild cognitive impairment (National Academies of Sciences et al., 2021a, 2021b). These clinical tendencies underscore the importance of longitudinal survey research such as the National Social Life, Health, and Aging Project (NSHAP) to gather data that accurately detect changes in cognition as Hispanics age. However, these trends also intimate the challenges in performing such research, especially among Hispanic Spanish speakers.

The Survey-Adapted Montreal Cognitive Assessment

Round 1 of NSHAP used the Short Portable Mental Status Questionnaire (SPMSQ), a 10-question cognitive screening measure originally designed to identify "organic brain deficits" (Pfeiffer, 1975), but this instrument was found to have a "ceiling effect"—96% of participants were considered to have "normal" cognitive functioning (Shega et al., 2014). A search for a more sensitive instrument led to the Montreal Cognitive Assessment (MoCA), a 28-item screening tool that evaluates multiple domains of cognitive function, including attention, memory, language, visuospatial abilities, and executive function, and that can be used to distinguish changes in normal aging related to mild cognitive impairment and early dementia among community-dwelling older adults in a variety of clinical settings (Nasreddine et al., 2005). The switch from the SPMSQ to the MoCA in NSHAP signified a crucial step toward accurately identifying cognitive deficits in the survey population; it permits differentiating normal aging from conditions like mild cognitive impairment and early dementia, thereby improving research precision.

The original 28-item MoCA, designed for clinical settings and clinician administration, was validated with Francophone and Anglophone participants in Montreal, Canada, and showed high test–retest reliability, good internal consistency, and equivalence in both languages (Nasreddine et al., 2005). In Round 2, NSHAP investigators aimed to adapt the MoCA for administration by nonmedically trained personnel (e.g., field interviewers) in a home setting, resulting in the development and validation of the MoCA-SA (English version) through revisions to its instructions and format (Kotwal et al., 2015; Shega et al., 2014). Later, a model was developed to estimate full MoCA scores from the MoCA-SA, which was subsequently validated with a clinical sample (Kotwal et al., 2015).

The MoCA-SA represents a tailored adaptation to suit the specific requirements of the NSHAP survey context. To start, researchers reduced administration times by 26% through strategic adjustments in question wording and item sequence, informed by cognitive interviews (Willis, 2005) conducted with a sample of older adults. These modifications were made to ensure that field interviewers could efficiently conduct in-home assessments, leading to higher completion rates (Kotwal et al., 2015; Shega et al., 2014).

Furthermore, researchers transformed the original one-page, paper-and-pen format into a Computer Assisted Personal Interview (CAPI) system, accompanied by tailored instructions for in-home administration. To address potential scoring issues in a survey context, a different approach was employed compared to clinical settings. Rather than simultaneously documenting and scoring responses, field interviewers were trained to accurately record task-related data during the in-home session, such as to write down every word beginning with the letter "F" that respondents could come up with in a given timeframe. Subsequently, dedicated scorers, trained specifically for item scoring, tallied the responses post-interview.

These adjustments, informed by empirical evidence and the unique survey context, resulted in an instrument that could be reliably administered through CAPI while retaining its capacity to assess a broad spectrum of cognitive abilities and ensure consistent performance across different demographic subgroups. Following these refinements, the final product was sent to a professional translation company to create a Spanish version that would be administered to Hispanic Spanish speakers, thus introducing measurement error for this sample.

Translation Problems

The translation of the MoCA-SA into Spanish was intended to enhance accessibility for Hispanic respondents. However, the Spanish version of the MoCA-SA did not undergo the same rigorous validation process as the original instrument. Specifically, some English words and phrasing in the original MoCA were substituted with new terms in the MoCA-SA, and these items were directly translated into Spanish without the benefit of cognitive interviews to verify their appropriateness for Spanish-speaking older adults. This oversight resulted in impactful changes in items that deviated from the validated, unabridged version of the Spanish MoCA (Delgado et al., 2019), potentially introducing measurement error. For instance, in the original Spanish MoCA, memory words were listed as follows: *rostro* (face), *seda* (silk), *iglesia* (church), *clavel* (carnation), and *rojo* (red). The MoCA-SA retained only two—*iglesia* (church) and *rojo* (red)—in the original English and Spanish. The words "silk" and "carnation" were substituted for "velvet" and "daisy" from the original English version because these words resonate better with English speakers in the United States. However, the translation of these words *terciopelo* (velvet) and *margarita* (daisy) may not necessarily resonate with Spanish-speaking populations residing in the United States. It is worth noting that while the word "face" was retained for the MoCA-SA, a different Spanish translation was used. The choice of the new term "cara," which is synonymous with "rostro," appears to be influenced by the translation vendor's preference rather than an alignment with preferences of Hispanic Spanish speakers in the United States, as there were no cognitive interviews conducted.

These subtle changes introduced potential measurement variability that could contribute to score disparities between English and Spanish speakers. Additionally, other deviations from the original MoCA (English version) included adding an example in the instructions for abstraction items (e.g., how are a banana and orange alike … they are both fruits). While data from cognitive interviews support these changes for English speakers, their suitability for Hispanic Spanish speakers is unknown due to the absence of equivalent interview data.

To explore potential measurement variability, we conducted a comparison analysis of MoCA-SA scores by ethnoracial and language groups (see Figure 1). Here and throughout the article, we employ the term "ethnoracial" to describe a socially constructed, perpetuated, and reinforced group identity that encompasses both racial and ethnocultural characteristics (Bean, 2018; Jiménez et al., 2015). While "Hispanic" is not a race, it is often treated as a racial category (Bean, 2018). In a similar vein, racial identities among Spanish speakers are complex. Recent census data indicate that many Spanish speakers (65%) identify as "White"; approximately one-third (32%) select "some other race"; and a minority identify as "Black" (2.5%), "Asian" (0.2%), or "White, non-Hispanic" (4.6%; Dietrich & Hernandez, 2022). In this study, we did not differentiate between "White Hispanic" and "Black Hispanic" due to data limitations. Additionally, we could not account for specific backgrounds related to cultural heritages. Instead, we used language preference as a proxy to capture one facet of Hispanic diversity, recognizing its significant impact on individuals' life trajectories and health outcomes.

Our initial analyses revealed that the MoCA-SA scores among NSHAP Hispanic participants in Rounds 2 and 3 were significantly lower compared to all other groups. Figure 1 presents a bar graph illustrating MoCA-SA mean scores for global cognition by ethnoracial group and language preference in Rounds 2 and 3. In Round 2, the average global score for Hispanic Spanish speakers was 17, a significant difference from their English-speaking counterparts across ethnoracial groups. In Round 3, average global cognition scores increased to 19 for Hispanic Spanish speakers, but they remained notably lower than their English-speaking counterparts across ethnoracial groups (see Table 1, Panel A, for additional descriptive details for Round 3). In both rounds, MoCA-SA scores for Hispanic English speakers were comparable to those of Blacks. Thus, in this study, we investigate whether statistical modifications can reduce measurement variance in global cognition scores among the two distinct language groups.

Drawing inspiration from the techniques employed by Iveniuk et al. in their work on mitigating measurement bias among English-speaking ethnoracial groups in relation to this issue, we applied modification indexes, two-parameter item response theory (IRT) models, and split-sample testing to pinpoint items that exhibit resilience to language-related bias among our Hispanic sample. Our analysis included an assessment of the predictive validity of both the complete MoCA-SA (18-item) and a shortened version, using self-reported dementia diagnoses, instrumental activities of daily living (IADLs), proxy reports of dementia-related deaths in Round 3, and dementia-related causes of death as documented in the National Death Index (NDI).

Method

Data

National Social Life, Health, and Aging Project is a nationally representative, longitudinal survey of community-dwelling

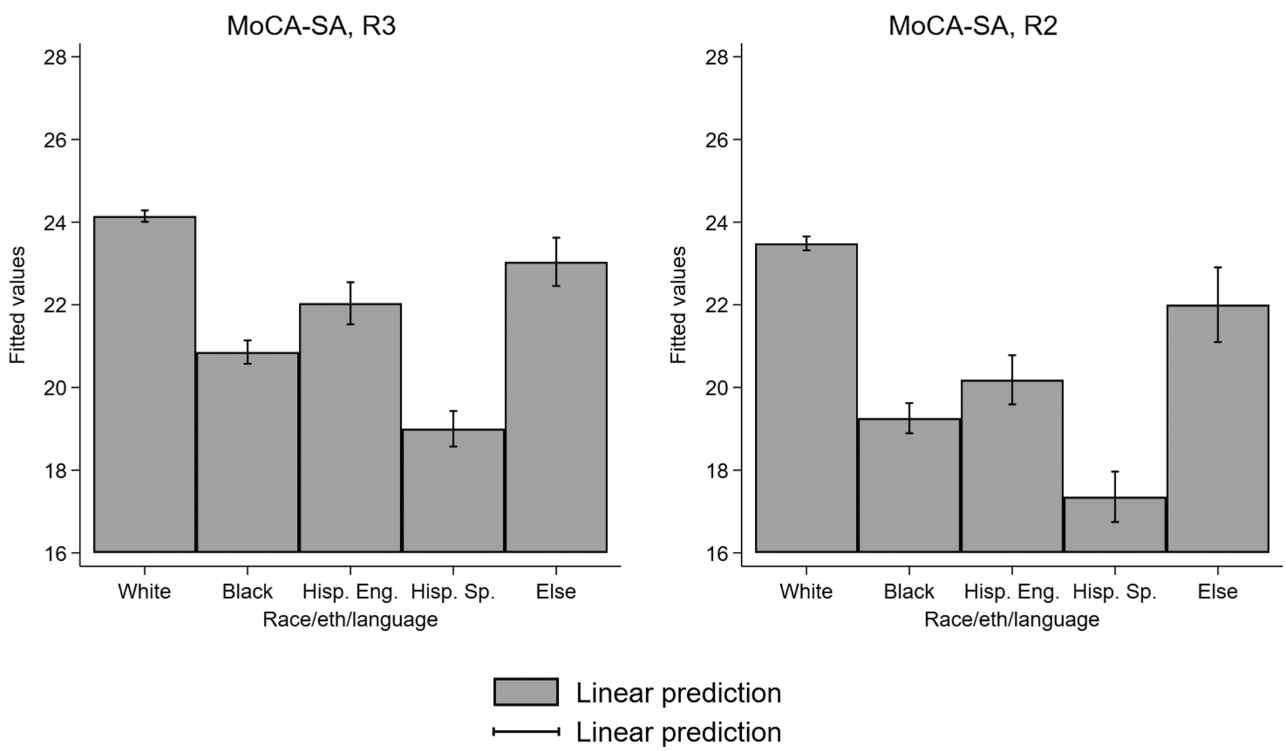

Figure 1. Bar graph of Montreal Cognitive Assessment-Survey Adaptation (MoCA-SA) scores by ethnoracial language groups, Rounds 2 and 3 (fitted values include 95% confidence intervals). Hisp. Eng. = Hispanic English; Hisp. Sp. = Hispanic Spanish.

Table 1. Descriptive Statistics for Sample

Variables	White		Black		Hispanic English		Hispanic Spanish		Else		Test of significance[a]
	Mean/%	*SD*/count	Mean/%	*SD*/count	Mean/%	*SD*/count	Mean/%	*SD*/count	Mean/%	*SD*/count	
Panel A: Round 3											
MoCA-SA	24.14	3.8	20.86	4.38	22.04	4.46	19	4.75	23.04	4.08	$p < .001$
MoCA-SAA	3.4	0.74	3.06	0.84	3.15	0.85	2.93	0.92	3.31	0.72	$p < .001$
Age	67.75	11.07	65.12	11.8	64.42	10.98	64.82	12.18	62.65	11.06	$p < .001$
Years of education	14.15	3.17	12.86	3.42	12.68	3.83	7.61	4.82	14.92	4.24	$p < .001$
Gender											$p = .116$
Male	45.28%	1,462	40.51%	316	47.92%	115	45.40%	153	42.31%	77	
Panel B: Round 2											
MoCA-SA	23.48	4.04	19.26	4.69	20.19	4.43	17.35	4.5	22	4.18	$p < .001$
MoCA-SAA	3.29	0.79	2.8	0.92	2.89	0.87	2.63	1.04	3.14	0.85	$p < .001$
Age	72.66	8.08	71.89	8.01	70.63	8.58	71.93	7.83	72.08	7.86	$p = .006$
Years of education	13.78	3.17	12.22	3.87	11.33	4.58	6.09	4.71	14.3	4	$p < .001$
Gender											$p = .504$
Male	45.94%	1,104	42.16%	215	45.83%	88	48.35%	88	48.19%	40	
Female	54.06%	1,299	57.84%	295	54.17%	104	51.65%	94	51.81%	43	

Notes: MOCA-SA = Montreal Cognitive Assessment-Survey Adaptation; MOCA-SAA = Montreal Cognitive Assessment-Survey Adaptation Abbreviated; *SD* = standard deviation.
[a]ANOVA test for continuous variables, chi-squared test for gender.

older adults in the United States. It commenced in 2005–2006 with Round 1, followed by interviews conducted every 5 years. The NSHAP sample was created using a multistage area probability sample design, intentionally oversampling Black and Hispanic adults to ensure a balanced representation (O'Muircheartaigh et al., 2009). In Round 1, the overall weighted response rate was 75.5%, resulting in 3,005 completed interviews. Interestingly, there was only a 4.5% difference in response rates between minority and nonminority households. Approximately 6% of interviews were conducted in Spanish (O'Muircheartaigh et al., 2009, 2014). For Round 2, conducted in 2010–2011, NSHAP reinterviewed the original respondents and expanded the sample to include the spouses and cohabiting partners of these respondents. They also attempted to interview Round 1 nonrespondents. The unconditional response rate in Round 2 was 73.9%, very close to Round 1's response rate of 75.5%, indicating minimal attrition, although the response rate was lower among older age cohorts (O'Muircheartaigh et al., 2014). In Round 3, which took place in 2015–2016, NSHAP conducted a third set of interviews with surviving respondents (Cohort 1) and introduced a new cohort of Baby Boomers and their spouses/partners (Cohort 2). The conditional response rate for Cohort 1 was 91%, while the response rate for Cohort 2 was 76% (O'Muircheartaigh et al., 2021).

The MoCA-SA was included in Rounds 2 and 3. Additional details about recruitment and sample characteristics can be found elsewhere (Lindau et al., 2007; O'Muircheartaigh et al., 2014; Waite et al., 2021). The data for this paper are drawn from Rounds 2 and 3, samples of 3,377 and 4,777, respectively. Given our focus on cognition in older adults, our analyses were restricted to respondents aged 50 and older (N = 4,768).

Our analytic sample comprises five ethnoracial and language groups: Black (n = 780; 16.4%), Hispanic English speakers (n = 240; 5%), Hispanic Spanish speakers (n = 337; 7.1%), Else (n = 182; 3.8%), and White (n = 3,229; 67.7%). Individuals in the "Else" grouping identified with an ethnoracial category other than Black, Hispanic, or White. Information on the cultural heritages of the Hispanic sample, as well as those in the other ethnoracial groups, was unavailable. The grouping by language reflects participants' preference for completing the NSHAP survey. Spanish was the sole alternative language option. We have no way to discern participants' fluency in another language, including English for those who took the Spanish survey. Participants in the Black, Else, and White groups completed the survey exclusively in English.

Table 1, Panels A and B, displays descriptive statistics for the Rounds 2 and 3 samples, organized by ethnoracial and language groups, with a focus on age, years of education, and gender.

The most significant disparity among these groups is in the realm of education, where Spanish-speaking Hispanics have nearly half the years of education compared to their English-speaking Hispanic counterparts. On average, in Round 3, the highest educational attainment was observed in the White group (14.2 years, standard deviation [SD] = 3.2) and the Else group (14.9 years, SD = 4.2). In contrast, the Black and Hispanic English speaker groups held similar intermediate levels of education (12.9 years, SD = 3.4; and 12.7 years, SD = 3.8, respectively), while the lowest levels of education were found among Hispanic Spanish speakers (7.6 years, SD = 4.8). The mean age was similar across groups, with the Else group being slightly younger compared to the others. The gender proportions indicate slightly more women compared to men across ethnoracial and language groups, which is expected. Similar trends can be observed in the Round 2 sample (see Table 1, Panel B). Furthermore, while there is measurement variance in cognitive performance across all groups for the MoCA-SA, this variance is most pronounced among Hispanic Spanish speakers. This pattern of variance is also noticeable in the abbreviated MoCA-SA, which we refer to as the MoCA-SAA, albeit to a considerably lesser degree than in the MoCA-SA.

Independent Variable: Cognition

Montreal Cognitive Assessment-Survey Adaptation

The MoCA-SA contains 18 items that assess global cognitive functioning (0–20; higher scores indicate better cognition) along eight domains: *orientation, naming, visual spatial, executive functioning, memory, attention, language,* and *abstraction* (Kotwal et al., 2015; Shega et al., 2014). Table 2 displays the domain, questions asked, and items represented in the queries. Except for the Serial 7 item listed in the *attention* domain, accurate responses were assigned a score of 1; incorrect answers received a 0. The Serial 7 item asked respondents to start with 100, subtract 7, and continue subtracting by 7 for up to 6 responses. This item was scored based on the number of correct responses, ranging 0–3. A score of 0 or 1 was assigned corresponding with the number of right answers. Two or three correct answers received a score of 2; four or five right answers were given a score of 3. There are two ways of scoring the NSHAP MoCA-SA: one includes a correction for educational level (i.e., essentially granting a point for people with low education), and one omits this correction (Shega et al., 2014). We use the version *without* this correction for education because an adjustment for education tends to apply to clinical samples (Adana Díaz et al., 2021). Because we are modifying the MoCA-SA for use with a probabilistic nonclinical sample and we control for education, the correction was unwarranted (Kotwal et al., 2015; Shega et al., 2014).

Dependent Variables: Four Consequential Outcomes

To assess the predictive performance of the MoCA-SAA, we incorporated several dependent variables, including self-reported dementia diagnoses, IADLs, proxy reports of dementia-related deaths in Round 3, and dementia-related causes of death documented in the NDI. We acknowledge that some of these outcomes, particularly those related to dementia identification (through self-reports, proxy reports, or death records), are relatively infrequent within the NSHAP sample. We calculated the following occurrence rates: self-reported dementia (2.88%; 68 out of 2,360), proxy reports of dementia-related deaths (1.92%; 65 out of 3,377), and dementia-related deaths recorded in the NDI (1.57%; 53 out of 3,377). This rarity can introduce bias in estimates and may pose challenges for model convergence (King & Zeng, 2001). To address this issue, we employed Firth's penalized maximum likelihood estimation (PMLE) function (Firth, 1993), which is further detailed in our data analysis section.

Self-reported dementia

Participants were asked, "Since your last interview (month/year), has a doctor told you that you have dementia, including

Table 2. Modification Indexes for MoCA-SA Items

Domain/question	ITEM	Black		Hispanic English		Hispanic Spanish		Else		Sum
		Int	Slope	Int	Slope	Int	Slope	Int	Slope	
Orientation: Give today's month and date.	MONTH									0
	DATE									0
Naming: Ask to identify a picture of a rhinoceros.	RHINO							30.71		30.71
Visuo-construction: Draw a clock and set the time to 10 after 11.	CONTOUR									0
	NUMBERS									0
	HANDS									0
Executive function: Draw a line going from a number to a letter in increasing order.	TRAILS	13.98				4.173		5.607		23.76
Memory: Repeat 5 words after a delay.	FACE			11.22				9.02		20.24
	VELVET									0
	CHURCH	3.975								3.975
	DAISY	5.887		8.123						14.01
	RED									0
Attention: Repeat 5 digits in forward order.	FORWARD	36.31	6.256	7.766	6.886	10.78	19.82			87.81
Repeat 3 digits in backward order.	BACKWARD									0
Subtract 7 from 100 and keep counting down by 7.	SERIAL 7									0
Language: List as many words starting with the letter F.	FLUENCY	6.866				5.566	5.057	5.136		22.63
Repeat a sentence.	CAT			10.72				28.46		39.18
Abstraction: How is a watch and a ruler alike?	WATCH/RULER							7.415		7.415

Note: MOCA-SA = Montreal Cognitive Assessment-Survey Adaptation.

Alzheimer's disease, vascular dementia, mixed dementia, or mild cognitive impairment?" (YES/NO).

Self-reported functional health problems

We used the IADL scale used in NSHAP to measure functional ability. We included items related to meal preparation, money management, shopping, light housework, medication administration, telephone use, driving during the day, and driving at night. Participants rated the degree of difficulty (0–3; 0 = no difficulty, 1 = some difficulty, 2 = much difficulty, and 3 = unable to do) for each item. We added the scores, then took the average; total scores ranged from 0 to 3.

Dementia as a contributing cause of death (proxy reports)

Mortality status at Round 3 was determined through proxy interviews. In these interviews, questionnaires were administered to a proxy person if the NSHAP respondent was deceased or too ill to participate in the data collection. If the person had died, the proxy was asked, "What was the main cause of [Respondent's NAME]'s death?" Possible response options were as follows: cancer (type specified), heart condition/heart attack, diabetes mellitus/high blood sugar, respiratory disease, stroke/cerebrovascular condition, dementia (such as Alzheimer's disease), accident, and other (cause specified). Most proxies were either a spouse (32%) or child/stepchild (37%). We coded cases as 1 if the proxy identified dementia (such as Alzheimer's disease) as a cause of death; all others were coded as 0.

Dementia as a main cause of death

NSHAP data were linked to the NDI and were used to ascertain cause of death for respondents who were deceased at Round 3. Cases were coded as 1 if the NDI identified a neurodegenerative disease—vascular disease or unspecified dementia, Alzheimer's disease, or other neurodestructive diseases—as a cause of death. All other cases were coded as 0 (including those who were still alive). The International Classification of Diseases codes used for the NDI measure were F01, F03, G30, and G31; all subcodes were used as well.

Age, sex (*female; male*), language *(English; Spanish*), and education were included as covariates. Education was reported as *years of schooling*, and age was reported in *years*. We also included ethnoracial grouping (*Black; Hispanic; White; Else*).

Data Analysis

We adopted a three-stage approach to develop the MoCA-SAA, and then compared it to the full version. We used both measures (i.e., the MoCA-SA and the MoCA-SAA) to predict four significant outcomes and conducted a comparative analysis of the results to assess the performance and utility of the abbreviated version.

Stage 1

We examined disparities in global MoCA-SA scores among different subgroups, including English-speaking individuals (Black, Hispanic, and White) and Spanish-speaking respondents (Hispanic only). Additionally, we conducted a thorough analysis of measurement invariance for each ethnoracial and language group to pinpoint the most significant variations among the test items and determine the extent to which these disparities correlated with specific groupings. Our primary hypothesis centered on the expectation of observing the most significant score variances between Hispanic Spanish speakers and English speakers, as depicted in Figure 1. Specifically, we anticipated that White respondents' scores would exhibit the greatest dissimilarity when compared to Hispanic Spanish language respondents when considering various ethnoracial groups. Conversely, we expected fewer score differences between Black and Hispanic English speakers.

These expectations are rooted in the distinctions and commonalities of two pivotal factors: language preference and educational attainment. While one's level of education is widely recognized as affecting one's access to better social goods, such as health benefits (Raghupathi & Raghupathi, 2020), limited English proficiency can also affect people's experiences and access to institutions and social opportunities in American society (Showstack, 2019; Terui & Hsieh, 2022; Zhao et al., 2021). Furthermore, both language preference and education can serve as indicators of exposure to broader American culture and society, as noted by Bean (2018). Additionally, the influence of low educational levels on MoCA scores has been consistently documented in prior research (Adana Díaz et al., 2021; Delgado et al., 2019; Milani et al., 2018). Across all ethnoracial groups in our sample, Hispanic Spanish speakers tend to have lower levels of education compared to their English-speaking counterparts. Consequently, we posited that fewer score differences would emerge between Black and Hispanic English speakers, who share the same survey language preference and possess similar levels of education.

Our analysis process began with the establishment of configural invariance, achieved through exploratory factor analyses conducted separately for each ethnoracial language group. Afterward, we proceeded to verify metric and scalar invariance by relaxing the assumption of equal loadings and thresholds. We supplemented these tests with log-likelihood tests for all possible comparisons, including configural versus metric, configural versus scalar, and scalar versus metric invariance, as per the methodology described by Putnick and Bornstein (2016). In our final set of analyses, we aimed to create the MoCA-SAA. The goal was to create a measurement tool with reduced variation in its performance across diverse ethnoracial and language groups, thereby improving its ability to be compared across these groups.

Stage 2

Modification indexes were generated to improve the model fit across ethnoracial language groups (Whittaker, 2012) by identifying items that would improve fit based on varied loadings and thresholds across groups. To calculate these indexes, the thresholds and loadings were controlled across groups (constrained as equal), and then, the parameter of interest was varied across groups. Next, we used chi-squared tests to compare the model-based covariance matrix to the actual covariance matrix for both the constrained and unconstrained versions of the model. The modification index is the difference between the two chi-squared tests, bestowing an improved model fit if the parameter varied across groups. Higher values indicate greater parameter variability across groups, an undesirable feature. We selected a cutoff of 3.84 for whether a parameter, and its corresponding item, was contributing to poor model fit and adding to measurement variance across groups. This threshold choice of 3.84 ($p < .05$) is congruent for a chi-squared test with one degree of freedom that compares two models with an extra-varying parameter for a specific group. All indicators were treated as dichotomous for this analysis, including the one ordinal measure of Serial 7, which was dichotomized at four or more correct answers.

Stage 3

We employed IRT models to further refine our selected measures, an essential step to maximize sensitivity to measurement invariance across different groups. IRT models organize questions within respondents and assign a "difficulty" parameter to each question, as well as an "ability" parameter to individuals, both of which account for variations in test-taking ability. This approach allows us to discern differences in how challenging each question is and how effectively it distinguishes between respondents with similar abilities (Nguyen et al., 2014).

For the selected items under consideration, we fit models that initially allow both the difficulty and discrimination parameters to vary by ethnoracial group. We then compared this model to one where neither parameter varies by group, using a standard likelihood ratio test. If the model with varying parameters indicated improved fit, we concluded that the item behaves differently among different ethnoracial groups. We also conducted a likelihood ratio test where only the difficulty parameter varied by group, while the discrimination parameter remained constrained by group. This enabled us to pinpoint the specific sources of variation in item performance.

Once we arrived at a model where the items that should vary by group did so, we proceeded with split-sample testing. In this step, we applied the established model to one half of the sample (randomly selected) and assessed how well its parameters fit the second half, assuming that if the model performs, we can reasonably conclude that the model's fit is not a result of chance.

In our final step, we constructed item-characteristic curves (ICCs) for each item that displayed variation across different ethnoracial groups. These ICCs provide valuable insights into the likelihood of successfully answering each item (on the y-axis) in relation to an individual's latent test-taking ability (on the x-axis), representing cognitive function in this context. Steeper curves on the ICCs indicate strong discrimination, signifying that once an individual reaches a certain level of ability, they can consistently answer the item correctly (Nguyen et al., 2014).

Following a thorough analysis of these rigorous tests, we derived an abbreviated version of the MoCA-SA measure, which we refer to as MoCA-SAA. This abbreviated measure includes only the items that have proven to be suitable and robust through our comprehensive analysis.

MoCA-SA and MoCA-SAA comparisons

We conducted a comparison between the MoCA-SAA and MoCA-SA measures. To assess each measure's performance, we examined the percentage of variance explained by language alone (English and Spanish) and by ethnoracial language groups. Additionally, we calculated Pearson correlations between the measures and test–retest Pearson correlations between Rounds 2 and 3, both for all groups pooled together and separately for each group. Finally, we evaluated the predictive validity of each measure for four significant outcomes: self-reported dementia diagnoses, IADLs, proxy reports of dementia-related deaths in Round 3, and dementia-related causes of death documented in the NDI.

When considered collectively, these consequential outcomes generally align. We found a moderately positive relationship between IADLs and respondent-reported dementia at Round 3, with a correlation coefficient of 0.27. Likewise, there was a correlation of 0.21 between proxy-reported dementia death and NDI-reported dementia death. We did not attempt to correlate the other outcomes because individuals who passed away between Rounds 2 and 3 would not have data available for the Round 3 outcomes.

Interrogating consequential outcomes

To address the fact that some of our outcomes, particularly those related to dementia identification (either through self- or proxy-report or by death records) are relatively rare events in the NSHAP sample, we employed Firth's PMLE function (Firth, 1993), which introduces a penalty factor into maximum likelihood estimates to correct for the expected bias arising from rare events, thus providing more accurate results.

For Firth's PMLE models, we used only standardized cognition scores. Functional health (as indicated by IADLs) was assessed using ordinary least squares regression while controlling for education, gender, and age. Lastly, we compared the predicted probabilities at the mean for each measure and in increments of 0.5 standard deviations (ranging from −1.5 to 1.5 standard deviations). All other predictor variables were held constant at their respective means when calculating predicted probabilities.

Exploratory factor analysis and all third-stage analyses were carried out in Stata v. 17, using the "factor," "regress," "firthlogit," and "margins" commands (StataCorp, 2021). Modification indexes and tests of measurement invariance were calculated in MPlus version 8 (Muthén & Muthén, 2017).

Results

We conducted an assessment to determine whether the MoCA-SAA could effectively mitigate measurement variance across different linguistic groups. Modification indexes were employed to enhance model fit across ethnoracial language groups by identifying items that would improve fit based on diverse loadings and thresholds. In Table 2, modification index results are shown, with cells left blank for indexes below the

Table 3. Pairwise Correlation by Language and Measure, Rounds 2 and 3 (Hispanic Only)

Language	Correlation between MoCA-SA and MoCA-SAA	Correlation between Rounds 2 and 3, MoCA-SA	Correlation between Rounds 2 and 3, MoCA-SAA
English	0.638	0.735	0.281
Spanish	0.617	0.684	0.409

Notes: MOCA-SA = Montreal Cognitive Assessment-Survey Adaptation; MOCA-SAA = Montreal Cognitive Assessment-Survey Adaptation Abbreviated.

3.84 threshold. The rightmost column summarizes modification index scores for each item. Shaded items are those with no modification indexes exceeding the 3.84 cutoff across all groups. Problematic items were then systematically removed based on their summed modification indexes (rightmost column, smallest to largest).

At each step, we assessed model fit improvements by conducting observed AIC, BIC, and log-likelihood tests upon removing each item. The AIC and BIC consistently decreased as each item was removed, while the log-likelihood test consistently produced significant results at $p < .001$. Consequently, we retained only the nine items with a zero-sum score as the final short form. When comparing our 9-item scale to the 7-item scale used for racial groups in Rounds 2 and 3, we found a substantial correlation of 0.91 (detailed analyses available upon request). Therefore, we proceeded to remove two additional items, "velvet" and Serial 7.

Lastly, we carried out IRT models with split-sample testing as described earlier, in which we applied the established model to one half of the sample (randomly selected) and assessed how well its parameters fit the second half (see Iveniuk et al., this volume, for additional technical details). This testing indicated only four items with difficulty parameters that were not statistically different across ethnoracial and language groups: reporting on today's date and the three items that make up the clock-drawing task. Therefore, while the MoCA-SA has 18 items, the MoCA-SAA has 4 items. For the sake of parsimony and ease of comparison among the ethnoracial and language groups, we recommend using the 4-item MoCA-SAA version.

We explored the extent of correlation between the two versions, both by language and across different time points. Table 3 presents the pairwise correlation table specifically for the Hispanic sample. In the case of Hispanic adults who speak either English or Spanish, we observed strong correlations between the MoCA-SAA version and the MoCA-SA (0.638 and 0.617, respectively). However, it is worth noting that within the Hispanic sample, the correlations between Rounds 2 and 3 were higher for the MoCA-SA (0.735 and 0.684, respectively) compared to the MoCA-SAA (0.281 and 0.409, respectively).

Table 4 presents the results concerning consequential outcomes, including self-reported dementia, IADL scores, proxy reports of dementia as a contributing cause of death, and NDI-recorded dementia as a cause of death, categorized by race and language group. Interaction terms between cognition and the two language variables were included for Hispanic adults (English and Spanish).

The findings in Table 4 reveal that, for all groups, both the long and short form measures exhibit similar behavior, although the cognition coefficients are smaller for the MoCA-SAA compared to the MoCA-SA. Furthermore, regardless of which form was used, we observed a more pronounced association between cognition and IADLs for Spanish-speaking Hispanics when compared to their English-speaking counterparts.

Comparing coefficients across models is potentially problematic in the case of logistic regressions because doing so assumes equal error variance for the key predictor (Allison, 1999). This challenge therefore affects our models of various consequential outcomes, for all but IADLs. To address this issue, we compare predicted values for each model. Figure 2 provides a visual representation of the predicted probabilities

Table 4. Predictive Coefficients of Cognition for the MoCA-SA and MoCA-SAA by Consequential Outcome

Variables	Self-reported dementia		IADL problems		Proxy-reported dementia death		NDI	
	MoCA-SA	MoCA-SAA	MoCA-SA	MoCA-SAA	MoCA-SA	MoCA-SAA	MoCA-SA	MoCA-SAA
Cognition	−1.23***, [a]	−0.60***	−0.17***, [a]	−0.13***	−1.16***	−0.70***	−1.26***, [a]	−0.75***
White	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.	Ref.
Black	−0.62	−0.07	−0.03	0.04	−0.34	0.25	−1.46**	−0.81
Hispanic English	0.28	0.07	−0.03	0.02	1.02	0.92	−1.83	−0.85
Hispanic Spanish	−0.92	−0.79	−0.19**	−0.05	0.23	0.45	1.30	0.48
Else	−1.20	−1.25	−0.01	0.03	1.09*	1.32*	0.06	0.15
Cog × Hispanic English	0.50	0.13	0.04	0.06	0.89	0.50	−0.53	−0.24
Cog × Hispanic Spanish	−0.10	−0.19	−0.13***	−0.11**	0.14	0.07	1.18	0.50
Female	−0.29	−0.30	0.11***	0.09***	0.55*	0.50	0.05	0.05
Years of education	0.10**	0.01	0.00	−0.01**	0.14***	0.08*	0.11**	0.05
Age	0.04*	0.07***	0.02***	0.02***	0.06**	0.08***	0.13***	0.15***
Constant	−8.47***	−9.27***	−1.02***	−1.09***	−10.84***	−11.63***	−15.81***	−16.66***
N	2,329	2,329	2,340	2,340	3,330	3,330	3,330	3,330

Notes: IADL = instrumental activities of daily living; MOCA-SA = Montreal Cognitive Assessment-Survey Adaptation; MOCA-SAA = Montreal Cognitive Assessment-Survey Adaptation Abbreviated; NDI = National Death Index.
[a]Cognition coefficients different at $p < .05$. $^*p < .05$. $^{**}p < .01$. $^{***}p < .001$.

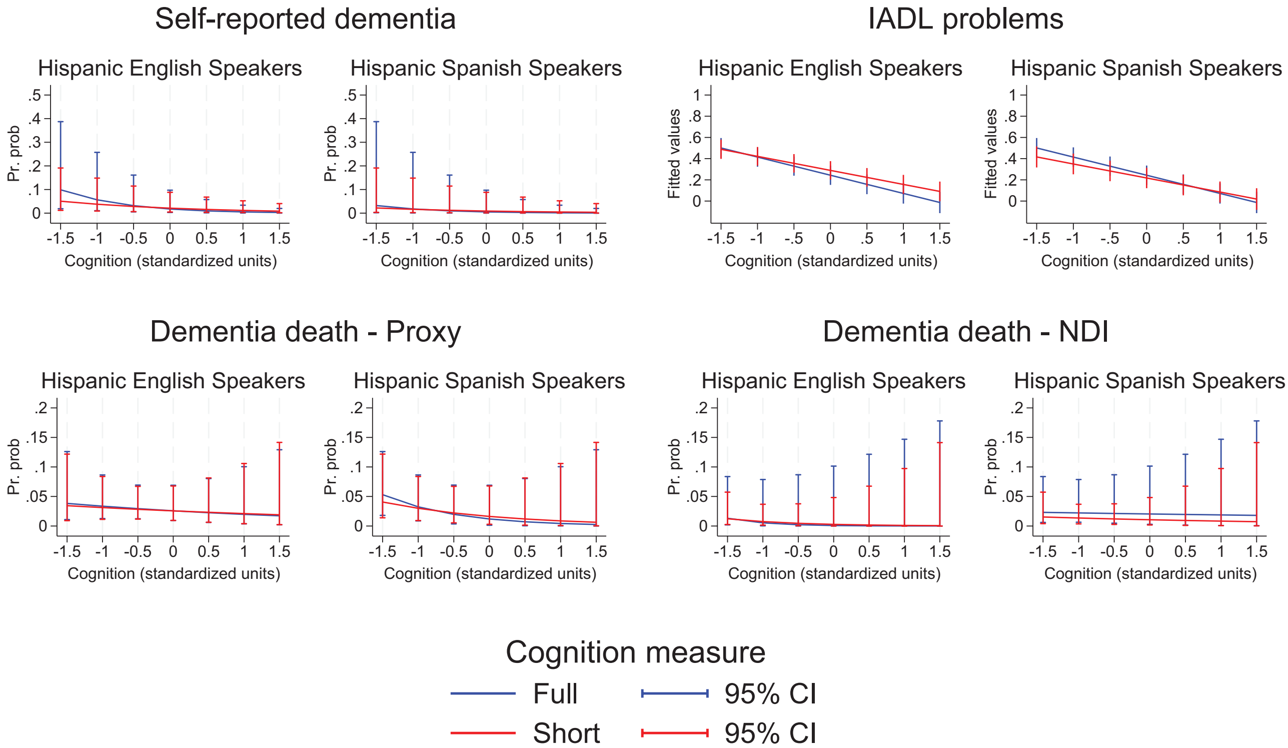

Figure 2. Predicted probabilities for each consequential outcome by language group and by MoCA versions (full—Montreal Cognitive Assessment-Survey Adaptation (MoCA-SA); short—Montreal Cognitive Assessment-Survey Adaptation Abbreviated (MoCA-SAA). CI = confidence interval; IADL = instrumental activities of daily living; NDI = National Death Index.

Table 5. Predictive Coefficients of Cognition for Both the Full and Shortened MoCA-SA

Variables	MoCA-SA	MoCA-SAA
White	0.42***	0.28***
Black	−0.28***	−0.12
Hispanic English	Ref.	Ref.
Hispanic Spanish	−0.23**	−0.01
Else	−0.05	0.03
Years of education	0.10***	0.05***
Female	0.08***	−0.06*
Age	−0.03***	−0.02***
Constant	0.28**	0.46***
N	4,574	4,574

Notes: IADL = instrumental activities of daily living; MOCA-SA = Montreal Cognitive Assessment-Survey Adaptation; MOCA-SAA = Montreal Cognitive Assessment-Survey Adaptation Abbreviated; NDI = National Death Index.
*$p < .05$. **$p < .01$. ***$p < .001$.

for each consequential outcome according to language group, using both the long and short forms.

For each of the consequential outcomes, the similarity in the slopes represented by the blue and red lines for both the full and short forms indicates similar predictive power for English- and Spanish-speaking Hispanics—with separation at low ends of the scale, indicating that the MoCA-SAA is potentially not as effective at capturing very poor cognitive ability.

Table 5 shows the associations between ethnoracial and language groups and cognition for both cognition measures. Compared to the MoCA-SA, we detected a weaker association for Spanish-speaking Hispanics and cognition than with the MoCA-SAA (−0.23 vs −0.01).

Figure 3 shows how the switch from MoCA-SA to MoCA-SAA affects cognitive gaps among ethnoracial groups for Rounds 2 and 3. The right-hand graphs display shallower columns compared to the left-hand graphs, indicating reduced measurement variability in the MoCA-SAA across all ethnoracial and language groups.

Discussion

The MoCA-SA represents a state-of-the-art instrument suitably adapted for a survey study. Nevertheless, researchers aiming to employ NSHAP's MoCA-SA for assessing global cognition and cognitive changes in Spanish-speaking older adults encounter translation and cultural congruency challenges, which raise legitimate concerns about potential measurement disparities based on language. While our findings do indicate statistical improvements, they do not entirely resolve the issues related to the translation of the MoCA-SA into Spanish.

Limitations

Several key limitations should be noted. Firstly, the MoCA-SAA is primarily designed to assess overall cognition and has not been validated to distinguish between various cognitive facets. Additionally, although our statistical adjustments offer a cross-cultural measure of overall cognition among specific ethnoracial language groups in the United States, its suitability for other ethnoracial or language groups, especially those speaking languages other than Spanish, remains uncertain. To extend its applicability to these groups, addressing the

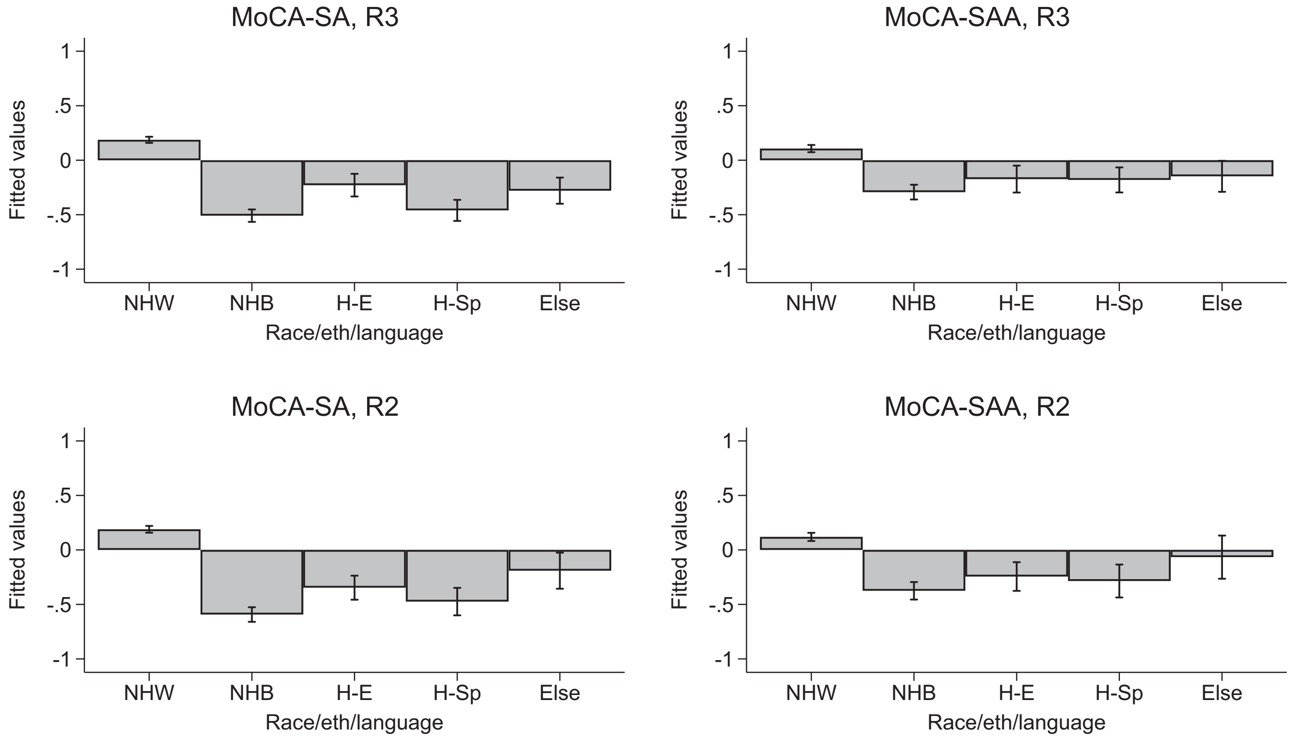

Figure 3. The effect on linear predications by ethnoracial group (relative to White respondents) for the Montreal Cognitive Assessment-Survey Adaptation (MoCA-SA) and Montreal Cognitive Assessment-Survey Adaptation Abbreviated (MoCA-SAA), Rounds 2 and 3.

translation challenges of the MoCA-SA and conducting cognitive interviews in the target language, as was done to validate the original MoCA for use in different countries, would be necessary (Cova et al., 2022; Delgado et al., 2019; Khan et al., 2022; O'Driscoll & Shaikh, 2017).

Secondarily, our data set lacked crucial information regarding participants' heritage backgrounds and places of birth. These factors represent significant aspects of heterogeneity within the pan-ethnic Hispanic population in the United States, which undoubtedly influence language use. Additionally, this diversity extends to various other dimensions, such as socioeconomic status (SES). However, we chose not to include SES in our analysis due to its potential multicollinearity with education. Furthermore, we did not account for the region of residency, given the sample size of English- and Spanish-speaking Hispanics in our study.

Another noteworthy limitation pertains to our inability to account for variations in English- and Spanish-speaking abilities among individuals within this demographic, primarily due to a lack of relevant information. Consequently, our study primarily focused on a single dimension of Hispanic diversity: those who opted for the Spanish version of the NSHAP survey and subsequently received the Spanish version of the MoCA-SA. Although we remain uncertain about their proficiency in English, we do know that they preferred Spanish when completing the NSHAP survey. It is plausible that Spanish-speaking Hispanic participants may possess varying degrees of English proficiency; some may have limited English skills, while others may speak English "well" or "very well" and still choose the Spanish survey option. Nevertheless, the decision to complete the extensive NSHAP survey in Spanish and, by extension, the MoCA-SA, introduces a specific form of measurement error.

By not considering these factors, our study may inadvertently oversimplify the analysis; neglecting other considerations such as how cultural backgrounds and linguistic dialects could introduce bias into the measure. Therefore, future research endeavors should aim to address these nuances by harmonizing independent translations that reflect diverse heritage backgrounds and by conducting cognitive interviews to capture cultural sensitivities and linguistic dialects that might influence how the instrument is interpreted.

Nonetheless, the current translated version of the MoCA-SA enables us to include Hispanic Spanish speakers who might otherwise be excluded, and the statistical adjustment outlined in this article offers a means to mitigate potential measurement disparities when assessing overall cognition. It represents the most viable option available at present.

Our pragmatic approach to reduce measurement variance in the MoCA-SA led to the discernment of four items that functioned well across groups, as indicated by the predicted probabilities of consequential outcomes and the reduction in coefficients between the two forms, as the shorter scale captures less variance. It is unsurprising that these items all reflected common everyday experiences such as reporting on today's date and the three items that make up the clock-drawing task. For example, knowing the current month or reading an analog clock tends to be universally understood, although the resonance of this latter task may change for future generations as digital clocks become increasingly ubiquitous.

In the absence of cognitive interviews, it is challenging to explain why other items varied across ethnoracial and language groups, but one problem is readily discernable. The straightforward translation of the MoCA-SA from English to Spanish imported language-specific biases. For example, the MoCA-SA uses the F-fluency question, which asks the person to list as many words as they can starting with the letter "F." The question works well in English but not in Spanish, where there are far fewer words, in particular nouns, that begin with "F." The validated Spanish version of the MoCA asks participants to list words that begin with the letter "P." These subtle differences can add considerable measurement noise, which can further compound variance attributed to lower levels of education.

Recommendations

We recommend using the shortened scale as a robustness check. However, for analysts who are particularly concerned about bias, alternative options exist, such as applying a partial measurement invariance model. In this model, the four bias-robust items are considered "anchors," permitting different intercepts and slopes across groups. However, this approach can be computationally demanding and challenging to execute correctly, leading to time and resource constraints. Therefore, in most cases and for most analysts, employing both the longer and shortened versions of the scale and comparing the outcomes provides a simpler and more practical approach.

Future Research

NSHAP's use of remote data collection methods in Round 4, such as web-based, paper and pen, and phone surveys, adds another layer of complexity to cognition assessment that merits exploration. Web-based methods require internet access and proficiency, which may exclude minority groups with lower rates of internet access and limited literacy or language skills required for these technologies. This concern was evident in the NSHAP-COVID study (Waite et al., 2022), which exclusively offered remote data collection options. Piedra et al. (2022) found that nearly 60% of White respondents in the NSHAP-COVID study preferred web-based data collection, while only 43% of Hispanic respondents did. They also observed that Hispanic NSHAP-COVID participants reported better self-rated health, higher education levels, and less material hardship compared to Hispanic participants in Round 3, suggesting that remote data collection may have introduced a selection bias (Piedra et al., 2022). These findings emphasize the crucial importance of meticulously evaluating the validity of measures when applied to diverse populations across varying data collection methods.

As remote data collection approaches continue to advance, the potential solutions to existing challenges may become more diverse. As these methods evolve, the application of refined statistical techniques, such as modification indexes and Firth PMLE models, will offer valuable means to address potential measurement differences. Nevertheless, when working with linguistic minority populations, rigorously translated instruments, complemented by cognitive interviews with individuals fluent in the target language, will significantly bolster our confidence that these tools accurately measure their intended constructs. It is imperative to have a comprehensive understanding of cognitive impairment within the expanding Hispanic population, and researchers must commit significant resources to ensure the development and availability of essential assessment tools.

Funding

This supplement is supported by the National Institute on Aging, National Institutes of Health (R01AG043538, R01AG048511) and NORC at the University of Chicago, which was responsible for the data collection. The content is solely the responsibility of the authors and does not necessarily represent the official views of NIA/NIH or NORC.

Conflict of Interest

None.

References

Adana Díaz, L., Arango, A., Parra, C., Rodríguez-Lorenzana, A., & Yacelga-Ponce, T. (2021). Impact of educational level on versions (basic and complete) of the Montreal Cognitive Assessment. *Dementia and Geriatric Cognitive Disorders*, 50(4), 341–348. https://doi.org/10.1159/000518747

Administration for Community Living. (2020). *Profile of Hispanic Americans age 65 and older: 2020*. U.S. Department of Health and Human Services. https://acl.gov/sites/default/files/Profile%20of%20OA/HispanicProfileReport2021.pdf

Allison, P. D. (1999). Comparing logit and probit coefficients across groups. *Sociological Methods & Research*, 28(2), 186–208. https://doi.org/10.1177/0049124199028002003

Angel, J. L., & Rote, S. M. (2021). Dementia support among older Mexican Americans: Is adult day care a viable alternative? In J. L. Angel, M. López Ortega, & L. M. Gutierrez Robledo (Eds.), *Understanding the context of cognitive aging: Mexico and the United States* (pp. 227–245). Springer International Publishing. https://doi.org/10.1007/978-3-030-70119-2_12

Bauer, A., Chen, C., & Alegria, M. (2010). English language proficiency and mental health service use among Latino and Asian Americans with mental disorders. *Medical Care*, 48(12), 1097–1104. https://doi.org/10.1097/MLR.0b013e3181f80749

Bean, F. D. (2018). Growing U.S. ethnoracial diversity: A positive or negative societal dynamic? *ANNALS of the American Academy of Political and Social Science*, 677(1), 229–239. https://doi.org/10.1177/0002716218765415

Cova, I., Nicotra, A., Maestri, G., Canevelli, M., Pantoni, L., & Pomati, S. (2022). Translations and cultural adaptations of the Montreal Cognitive Assessment: A systematic and qualitative review. *Neurological Sciences*, 43(1), 113–124. https://doi.org/10.1007/s10072-021-05716-y

Delgado, C., Araneda, A., & Behrens, M. I. (2019). Validation of the Spanish-language version of the Montreal Cognitive Assessment test in adults older than 60 years (Validación del instrumento Montreal Cognitive Assessment en español en adultos mayores de 60 años). *Neurologia*, 34(6), 376–385. https://doi.org/10.1016/j.nrl.2017.01.013

Dietrich, S., & Hernandez, E. (2022). Language use in the United States: 2019. *American Community Survey Reports*. https://www.census.gov/content/dam/Census/library/publications/2022/acs/acs-50.pdf

Firth, D. (1993). Bias reduction of maximum likelihood estimates. *Biometrika*, 80(1), 27–38. https://doi.org/10.2307/2336755

Fitten, L. J., Ortiz, F., Fairbanks, L., Bartzokis, G., Lu, P., Klein, E., Coppola, G., & Ringman, J. (2014). Younger age of dementia diagnosis in a Hispanic population in southern California. *International Journal of Geriatric Psychiatry*, 29(6), 586–593. https://doi.org/10.1002/gps.4040

Ginde, A. A., Clark, S., & Camargo, C. A. (2009). Language barriers among patients in Boston emergency departments: Use of medical interpreters after passage of interpreter legislation. *Journal of Immigrant and Minority Health*, 11(6), 527–530. https://doi.org/10.1007/s10903-008-9188-5.

González, H. M., Tarraf, W., Schneiderman, N., Fornage, M., Vásquez, P. M., Zeng, D., Youngblood, M., Gallo, L. C., Daviglus, M. L., Lipton, R. B., Kaplan, R., Ramos, A. R., Lamar, M., Thomas, S., Chai, A., & DeCarli, C. (2019). Prevalence and correlates of mild cognitive impairment among diverse Hispanics/Latinos: Study of

Latinos-Investigation of Neurocognitive Aging results . *Alzheimer's & Dementia*, 15(12), 1507–1515. https://doi.org/10.1016/j.jalz.2019.08.202
Goth, U. S., & Strøm, B. S. (2018). Language disintegration: Communication ability in elderly immigrants with dementia. *Lancet Public Health*, 3(12), e563. https://doi.org/10.1016/S2468-2667(18)30184-1
Haan, M. N., Mungas, D. M., Gonzalez, H. M., Ortiz, T. A., Acharya, A., & Jagust, W. J. (2003). Prevalence of dementia in older Latinos: The influence of type 2 diabetes mellitus, stroke and genetic factors. *Journal of the American Geriatrics Society*, 51(2), 169–177. https://doi.org/10.1046/j.1532-5415.2003.51054.x
Jiménez, T. R., Fields, C. D., & Schachter, A. (2015). How ethnoraciality matters: Looking inside ethnoracial "groups." *Social Currents*, 2(2), 107–115. https://doi.org/10.1177/2329496515579765
Khan, G., Mirza, N., & Waheed, W. (2022). Developing guidelines for the translation and cultural adaptation of the Montreal Cognitive Assessment: Scoping review and qualitative synthesis. *BJPsych Open*, 8(1), e21. https://doi.org/10.1192/bjo.2021.1067
King, G., & Zeng, L. (2001). Logistic regression in rare events data. *Political Analysis*, 9(2), 137–163. https://doi.org/10.1093/oxfordjournals.pan.a004868
Kotwal, A., Schumm, L. P., Kern, D. W., McClintock, M., Waite, L., Shega, J. W., Huisingh-Scheetz, M., & Dale, W. (2015). Evaluation of a brief survey instrument for assessing subtle differences in cognitive function among older adults. *Alzheimer Disease & Associated Disorders*, 29(4), 317–324. https://doi.org/10.1097/wad.0000000000000068
Krogstad, J. M., & Gonzalez-Barrera, A. (2015). *A majority of English-speaking Hispanics in the U.S. are bilingual*. Pew Research Center. https://www.pewresearch.org/fact-tank/2015/03/24/a-majority-of-english-speaking-hispanics-in-the-u-s-are-bilingual/
Krogstad, J. M., Stepler, R., & Lopez, M. H. (2015). *English proficiency on the rise among Latinos: U.S. born driving language changes*. Pew Research Center. http://www.pewhispanic.org/2015/05/12/english-proficiency-on-the-rise-among-latinos/
Lindau, S. T., Schumm, L. P., Laumann, E. O., Levinson, W., O'Muircheartaigh, C. A., & Waite, L. J. (2007). A study of sexuality and health among older adults in the United States. *New England Journal of Medicine*, 357(8), 762–774. https://doi.org/10.1056/NEJMoa067423
Massett, H. A., Mitchell, A. K., Alley, L., Simoneau, E., Burke, P., Han, S. H., Gallop-Goodman, G., & McGowan, M. (2021). Facilitators, challenges, and messaging strategies for Hispanic/Latino populations participating in Alzheimer's disease and related dementias clinical research: A literature review. *Journal of Alzheimer's Disease*, 82(1), 107–127. https://doi.org/10.3233/JAD-201463. https://content.iospress.com/articles/journal-of-alzheimers-disease/jad201463
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 & Dementia*, 15(1), 17–24. https://doi.org/10.1016/j.jalz.2018.06.3063
McMurtray, A., Saito, E., & Nakamoto, B. (2009). Language preference and development of dementia among bilingual individuals. *Hawaii Medical Journal*, 68(9), 223–226. PMID: 19842364; PMCID: PMC4335728. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4335728/pdf/nihms660694.pdf
Mehta, K. M., & Yeo, G. W. (2017). Systematic review of dementia prevalence and incidence in United States race/ethnic populations . *Alzheimer's & Dementia*, 13(1), 72–83. https://doi.org/10.1016/j.jalz.2016.06.2360
Milani, S. A., Marsiske, M., Cottler, L. B., Chen, X., & Striley, C. W. (2018). Optimal cutoffs for the Montreal Cognitive Assessment vary by race and ethnicity. *Alzheimer's & Dementia (Amsterdam, Netherlands)*, 10, 773–781. https://doi.org/10.1016/j.dadm.2018.09.003
Muthén, L. K., & Muthén, B. O. (2017). *MPlus user's guide* (8th ed.). Muthén & Muthén.
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. https://doi.org/10.1111/j.1532-5415.2005.53221.x
National Academies of Sciences, Engineering, & Medicine. (2021a). *Meeting the challenge of caring for persons living with dementia and their care partners and caregivers: A way forward*. The National Academies Press. https://doi.org/10.17226/26026
National Academies of Sciences, Engineering, & Medicine. (2021b). *Reducing the impact of dementia in America: A decadal survey of the behavioral and social sciences*. The National Academies Press. https://doi.org/10.17226/26175
Nguyen, T. H., Han, H. -R., Kim, M. T., & Chan, K. S. (2014). An introduction to item response theory for patient-reported outcome measurement. *Patient*, 7(1), 23–35. https://doi.org/10.1007/s40271-013-0041-0
O'Driscoll, C., & Shaikh, M. (2017). Cross-cultural applicability of the Montreal Cognitive Assessment (MoCA): A systematic review. *Journal of Alzheimer's Disease*, 58(3), 789–801. https://doi.org/10.3233/JAD-161042
O'Muircheartaigh, C., Eckman, S., & Smith, S. (2009). Statistical design and estimation for the National Social Life, Health, and Aging Project. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 64(suppl_1), i12–i19. https://doi.org/10.1093/geronb/gbp045
O'Muircheartaigh, C., English, N., Pedlow, S., & Kwok, P. K. (2014). Sample design, sample augmentation, and estimation for Wave 2 of the NSHAP. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69(Suppl_2), S15–S26. https://doi.org/10.1093/geronb/gbu053
O'Muircheartaigh, C., English, N., Pedlow, S., & Schumm, L. P. (2021). Sample design and estimation in the National Social Life, Health, and Aging Project: Round 3 (2015–2016). *Journals of Gerontology, Series B: Psychological Sciences sand Social Sciences*, 76(Suppl_3), S207–S214. https://doi.org/10.1093/geronb/gbab182
Pfeiffer, E. (1975). A Short Portable Mental Status Questionnaire for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatrics Society*, 23(10), 433–441. https://doi.org/10.1111/j.1532-5415.1975.tb00927.x
Piedra, L. M., Andrade, F. C. D., & Larrison, C. R. (2011). Building response capacity: The need for universally available language services. In L. P. Buki & L. M. Piedra (Eds.), *Creating infrastructures for Latino mental health* (pp. 55–76). Springer Publishing Co. https://doi.org/10.1007/978-1-4419-9452-3_3
Piedra, L. M., Howe, M. J. K., Francis, J., Gutwein, M., & Montoya, Y. (2022). Latinos and the pandemic: Results from the National Social Life, Health, and Aging Project—COVID-19 study. *Journal of Applied Gerontology*, 41(5), 1465–1472. https://doi.org/10.1177/07334648211066920
Pot, A., Keijzer, M., & De Bot, K. (2020). The language barrier in migrant aging. *International Journal of Bilingual Education and Bilingualism*, 23(9), 1139–1157. https://doi.org/10.1080/13670050.2018.1435627
Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41(Sept.), 71–90. https://doi.org/10.1016/j.dr.2016.06.004
Raghupathi, V., & Raghupathi, W. (2020). The influence of education on health: An empirical assessment of OECD countries for the period 1995–2015. *Archives of Public Health*, 78(1), 20. https://doi.org/10.1186/s13690-020-00402-5
Schenker, Y., Wang, F., Selig, S. J., Ng, R., & Fernandez, A. (2007). The impact of language barriers on the documentation of informed consent at a hospital with on-site interpreter services. *Journal of General Internal Medicine*, 22(Suppl 2), 294–299. https://doi.org/10.1007/s11606-007-0359-1

Schyve, P. M. (2007). Language differences as a barrier to quality and safety in health care: The joint commission perspective. *Journal of General Internal Medicine*, 22(Suppl 2), 360–361. https://doi.org/10.1007/s11606-007-0365-3

Shega, J. W., Sunkara, P. D., Kotwal, A., Kern, D. W., Henning, S. L., McClintock, M. K., Schumm, P., Waite, L. J., & Dale, W. (2014). Measuring cognition: The Chicago Cognitive Function measure in the National Social Life, Health and Aging Project, wave 2. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 69(Suppl_2), S166–76. https://doi.org/10.1093/geronb/gbu106

Showstack, R. (2019). Patients don't have language barriers; the healthcare system does. *Emergency Medicine Journal*, 36(10), 580–581. https://doi.org/10.1136/emermed-2019-208929

StataCorp. (2021). *Stata statistical software: Release 17*. StataCorp LLC.

Terui, S., & Hsieh, E. (2022). Language barriers as a social phenomenon: Distinctive impacts on health communication in Japan and the United States. In D. K. D. Kim & Kreps, G. L. (Eds.), *Global health communication for immigrants and refugees: Cases, theories, and strategies* (1st ed., pp. 63–77). Routledge. https://doi.org/10.4324/9781003230243

Waite, L. J., Hawkley, L., Kotwal, A. A., O'Muircheartaigh, C., Schumm, L. P., & Wroblewski, K. (2021). Analyzing birth cohorts with the National Social Life, Health, and Aging Project. *Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*, 76(Suppl_3), S226–S237. https://doi.org/10.1093/geronb/gbab172

Waite, L. J., Cagney, K. A., Dale, W., Hawkley, L. C., Huang, E. S., Lauderdale, D. S., Laumann, E. O., McClintock, M. K., O'Muircheartaigh, C. A., & Schumm, L. P. (2022). *National Social Life, Health, and Aging Project (NSHAP): Round 3 and COVID-19 Study, [United States], 2015–2016, 2020–2021 Inter-university Consortium for Political and Social Research [distributor]*. https://doi.org/10.3886/ICPSR36873.v7

Whittaker, T. A. (2012). Using the modification index and standardized expected parameter change for model modification. *Journal of Experimental Education*, 80(1), 26–44. https://doi.org/10.1080/00220973.2010.531299

Willis, G. B. (2005). Cognitive interviewing in practice: Think-aloud, verbal probing, and other techniques. In G. B. Willis (Ed.), *Cognitive interviewing: A tool for improving questionnaire design* (pp. 42–65). Sage Publications. https://doi.org/10.4135/9781412983655.n4

Wu, S., Vega, W. A., Resendez, J., & Jin, H. (2016). *Latinos and Alzheimer's disease: New numbers behind the crisis*. USC Edward R. Roybal Institute on Aging and the LatinosAgainstAlzheimer's Network. https://www.usagainstalzheimers.org/sites/default/files/Latinos-and-AD_USC_UsA2-Impact-Report.pdf

Zhao, Y., Segalowitz, N., Voloshyn, A., Chamoux, E., & Ryder, A. G. (2021). Language barriers to healthcare for linguistic minorities: The case of second language-specific health communication anxiety. *Health Communication*, 36(3), 334–346. https://doi.org/10.1080/10410236.2019.1692488