



Published in final edited form as:

Alzheimer Dis Assoc Disord. 2015 ; 29(4): 317–324. doi:10.1097/WAD.0000000000000068.

Evaluation of a brief survey instrument for assessing subtle differences in cognitive function among older adults

Ashwin A. Kotwal, MD, MS^{1,2}, L Philip Schumm, MA², David W. Kern, MA³, Martha K. McClintock, PhD³, Linda J. Waite, PhD⁴, Joseph W. Shega, MD¹, Megan J. Huisingsh-Scheetz, MD¹, and William Dale, MD, PhD¹

¹University of Chicago, Department of Medicine, Section of Geriatrics & Palliative Medicine

²University of Chicago, Department of Health Studies

³University of Chicago, Department of Comparative Human Development and Institute for Mind and Biology

⁴University of Chicago, Department of Sociology

Abstract

Most measures of cognitive function used in large-scale surveys of older adults have limited ability to detect subtle differences across cognitive domains, and standard clinical instruments are impractical to administer in general surveys. The Montreal Cognitive Assessment (MoCA) can address this need, but has limitations in a survey context. Therefore, we developed a survey-adaptation of the MoCA, called the MoCA-SA, and describe its psychometric properties in a large national survey. Using a pretest sample of older adults (n=120), we reduced MoCA administration time by 26%, developed a model to accurately estimate full MoCA scores from the MoCA-SA, and tested the model in an independent clinical sample (n=93). The validated 18-item MoCA-SA was then administered to community-dwelling adults aged 62–91 as part of the National Social Life Health and Aging Project (NSHAP) Wave 2 sample (n=3,196). In NSHAP Wave 2, the MoCA-SA had good internal reliability (Cronbach $\alpha=0.76$). Using item-response models, survey-adapted items captured a broad range of cognitive abilities and functioned similarly across gender, education, and ethnic groups. Results demonstrate that the MoCA-SA can be administered reliably in a survey setting while preserving sensitivity to a broad range of cognitive abilities and similar performance across demographic subgroups.

Keywords

cognitive assessment; cognitive impairment; older adults; item response theory; survey; MoCA

Corresponding Author: Ashwin Kotwal, University of Chicago Medicine, Department of Medicine, Section of Geriatrics & Palliative Medicine, MC6098, 5841 S. Maryland Avenue Chicago, IL 60637, ashwin.kotwal@gmail.com, Phone: 773-834-0508 Fax: 773-702-3538.

Conflict of interest

The authors report no conflict of interests.

INTRODUCTION

A better understanding of early, often subclinical cognitive loss is important as it represents a potential target for behavioral or clinical interventions, and it could be linked to early functional impairment.¹ To date, however, early cognitive declines have been difficult to assess in population-based research, primarily because available survey measures were designed to detect more advanced cognitive impairment or to determine the competence of respondents to participate in national surveys. For example, the Short Portable Mental Status Questionnaire (SPMSQ) and Mini-Mental Status Exam (MMSE) can be completed in a survey format,^{2, 3} but both have experienced declining use due to their lack of sensitivity for detecting modest cognitive changes, especially in community samples.⁴⁻⁶ The Telephone Interview for Cognitive Status (TICS) has been integrated as part of the Health and Retirement Study,⁷ but it focuses on cognitive domains evaluable by phone assessment, such as language or memory, and omits other important determinants of cognition, particularly executive function and visuospatial skills.⁸⁻¹⁰ Alternatively, available clinical tools, such as full neuropsychological testing, are limited in survey settings due to high cost, lengthy administration time, and a need for medically-knowledgeable personnel to administer.¹¹

One candidate measure to address this need for detecting milder cognitive loss in a survey setting is the Montreal Cognitive Assessment (MoCA), a multidomain clinical tool developed to differentiate cognitive changes of normal aging from mild cognitive impairment (MCI) and early dementia.¹² However, several challenges exist in incorporating this clinical tool into a large survey setting without medical personnel. First, previous research has raised questions about the reliability of the MoCA in a survey context (Cronbach $\alpha = 0.50$ in survey vs 0.75 in a clinical sample).¹³ Consequently, it requires alteration, standardized scoring algorithms and use of standard administration protocols to improve clarity and ease the data collection process in a national home-based survey.¹⁴ Second, although the clinical administration time of the MoCA has been reported as <10 minutes,¹² the average administration time in an in-home survey was over 15 minutes.⁶ This increased time may prevent the use of the clinical MoCA in a typical time-constrained general survey which evaluates multiple domains of health and other areas. Third, while psychometric work has been done with the MoCA, the majority has been conducted on clinical samples or samples drawn from ethnically homogenous populations,¹⁵⁻¹⁸ providing minimal evidence for how the measure performs when administered by non-medical survey interviewers to the ethnically heterogeneous US national sample.

Given the potential of the MoCA, but these limitations in a survey context, we adapted the MoCA for use as part of a computer-assisted survey, with alterations to allow for reliable in-home administration by trained, non-medical personnel.⁶ In this study, we report on developing the survey-adapted measure, known as the MoCA-SA, in greater detail and evaluate its psychometric properties in three independent samples, including the National Social Life Health and Aging Project (NSHAP) Wave 2 sample. Our purpose is to contribute to prior psychometric work conducted on the MoCA. We have the following objectives: 1) to demonstrate how the MoCA-SA can accurately estimate full MoCA scores; 2) to examine the properties of the MoCA-SA items in a large US nationally-representative sample of

community-dwelling adults; and 3) to assess how the measure performs across key demographic subgroups.

METHODS

Data and sample

We used three separate data sources for our analyses: 1) NSHAP Wave 2 (W2) pretest sample, 2) NSHAP W2 age-eligible sample, and 3) an outpatient clinical sample of frail older adults. The NSHAP W2 pretest comprises a purposive sample of community-dwelling adults aged 46–89 (n=120), selected to reflect the composition of participants in NSHAP W2. Second, we used the nationally-representative NSHAP W2 data, collected between August 2010 and May 2011; this sample includes 3,196 community-dwelling respondents born between 1920–1947. NSHAP W2 was conducted by the National Opinion Research Center (NORC) using trained professional interviewers who administered the instrument using computer assisted personal interviewing (CAPI) in both English and Spanish. W2 had a weighted overall response rate of 76.9%. Individuals judged by the interviewer to be unable to complete the interview, either because of physical or cognitive limitations, were not interviewed. Further details on the NSHAP samples are available elsewhere.^{19, 20} Third, we used a dataset from the University of Chicago’s South Shore Senior Center which includes 93 ethnically diverse, frail older adults. Interviews with clinic patients were conducted by physicians and trained research assistants.

Instrument Development

A reformatted version of the MoCA for administration in a computer-based survey was developed by an interdisciplinary research team for inclusion into NSHAP W2. Based on an extensive literature review, the MoCA was chosen as a starting framework due to its assessment of multiple cognitive domains, ability to detect milder degrees of cognitive impairment, reliability, and validity in clinical settings. The MoCA includes 28 items representing 6 cognitive domains: “executive function,” “visuospatial skills,” “language,” “attention, concentration, and working memory,” “orientation,” and “short-term memory.”¹² In initial pilot testing, the MoCA was adapted for survey administration in conjunction with cognitive interviewing to a clinical sample of older adults. Based on results and feedback from participants, the format was altered in the following ways: 1) rewording of some questions and administration instructions for improved understanding, 2) reordering items to maximize completion rates, and 3) modifying the layout from one written page to a computer-based format compatible with CAPI technology. Furthermore, survey data quality and reproducibility were optimized by minimizing items scored by interviewers in the field, so that they were blindly scored later by trained personnel applying a standardized scoring protocol.

The reformatted version of the MoCA was administered to the NSHAP W2 Pretest sample. The pretest data was used to further refine administrative technique for CAPI and facilitate item selection to reduce administration time. Given the constraints of an omnibus survey in NSHAP W2, we aimed to reduce administration time from 15.6 minutes to under 12 minutes (i.e. at least a 23% reduction). To achieve this goal we set four criteria for selecting a subset

of items: 1) preservation of questions from each of the 6 cognitive domains, 2) preferential inclusion of difficult items in each domain to ensure discrimination of cognition in a relatively high-functioning population, 3) elimination of items that were difficult to administer in the field based on interviewer feedback on the pretest, and 4) high correlation between the shortened form and the full scale. After selection of items, we determine whether the shortened MoCA-SA can be reliably administered in a large-scale survey and examine its psychometric properties.

Demographic and health characteristics

We included several covariates in our analysis. Age is measured in years (continuous). Education is categorized as less than high school (HS), HS/GED, some college or vocational certification, or bachelors degree or more. Race/Ethnicity is defined as white, African American (AA), Hispanic non-AA, or other. Marital status is defined as married, divorced, widowed, or never married. To measure health status, we use self-rated health with possible responses of excellent, very good, good, fair or poor. The self-reported comorbidities measured included: dementia, heart disease, chronic obstructive pulmonary disease or asthma, arthritis, diabetes, and stroke. Participants reported their level of difficulty with Instrumental Activities of Daily Living (IADLs) (shopping, managing finances, light housework) and Activities of Daily Living (ADLs) (walking across the room, and bathing).

Analytic Approach

Relationship of full MoCA Scores to MoCA-SA scores—Using data from both the NSHAP W2 pretest and clinic samples, we did a regression of the full MoCA scores on MoCA-SA scores separately for each sample. Using pooled data and an interaction term, we tested the null hypothesis that the regression equation is the same in both samples. We report the resulting prediction equation in the pretest sample, and summarize the accuracy of the predictions using the 95% forecast intervals in each sample.²¹

Item-response model—We fit a generalization of the well-known Rasch model known as the “partial credit” model to the MoCA-SA items in each of the three datasets.^{22, 23} According to this model, the probability of respondent i scoring higher than response k on item j is written as

$$\text{logit}(\Pr(y_{ij} > k | \alpha_i)) = \alpha_i - \theta_{jk} \quad (1)$$

where y_{ij} is respondent i 's score for the j 'th item and θ_{jk} represents the “difficulty” of scoring higher than k on the j 'th item ($k = 1, 2, \dots, m - 1$ where m is the number of possible values for item j). Note that in the case of a binary item, $m = 2$, so θ_{jk} reduces to θ_j which is then simply the difficulty of the j 'th item. The parameters α_i represent the latent (unobserved) cognitive “abilities” of each respondent, arrayed along a single dimension. Like the Rasch model, Model (1) represents differences in cognitive ability on the same scale as differences in difficulty between items; thus, a unit increase in cognitive ability has the same effect on the probability of getting a given item correct as a unit decrease in item difficulty. To fit the model, we assume that the α_i are distributed $N(0, \sigma^2)$. The model was fit via maximum likelihood using the *gllamm* package in Stata v12.1.^{24, 25}

A limitation of Model (1) is that it assumes that a unit increase in ability has an equivalent effect (on the logit scale) for each item on the probability of answering the item correctly. This limitation may be addressed by extending Model (1) in the following way:

$$\text{logit}(\Pr(y_{ij} > k | \alpha_i)) = \lambda_j \alpha_i - \theta_{jk} \quad (2)$$

In this model, λ_j are referred to as “discrimination” parameters, allowing for variation across items in how informative they are with respect to differences in cognitive ability (these are similar to factor loadings for a factor analytic model). We fit Model (2) to the data from the NSHAP W2 sample, and performed a likelihood ratio test comparing Model (1) to Model (2).

For both item-response models, we obtained estimates of the α_i by using the posterior means from the fitted model. We calculated the correlation between the estimates of cognitive ability from the two models and the sum of the item scores from the MoCA-SA.

Differential Item Functioning—The model assumes that the relative difficulty of the items is the same across administration settings and for all individuals. We tested these assumptions in two ways. First, we fit Model (1) to a pooled dataset containing the data from all three samples, including interaction terms between sample type and each item difficulty. A likelihood ratio test was used to test the null hypothesis that the item difficulties are equivalent across the three samples, and Wald tests were used to test whether the difficulties of a specific item differed across the samples. Second, we fit the model to the NSHAP W2 sample including interaction terms between several demographic variables (sex, race/ethnicity and education) and the item difficulties; this analyses was conducted with a Rasch model in which the subtract 7s item was dichotomized as 0–1 points versus 2–3 points. For each demographic variable, a likelihood ratio test was performed to test whether the item difficulties varied across the subgroups defined by that variable. Results from these analyses are presented in tabular form and by plotting the profiles of estimated item difficulties by subgroup.

RESULTS

Sample characteristics

Characteristics of the NSHAP pretest, NSHAP W2, and clinical samples are presented in Supplementary Table 1. NSHAP pretest individuals are more educated, have higher self-rated health, and better functional status compared to the other samples. They have mean MoCA scores of 23.4 (Cronbach alpha=0.774). Individuals in NSHAP W2 are on average 73 years old, distributed evenly across the four education categories, and on average report being in good physical health and having good functional status. The clinical sample from a geriatrics clinic has a mean age of 84 years old, is 76.3% female, 67% African American, and has far more disabilities than the NSHAP samples. Additionally, 24.5% of clinic individuals have a known diagnosis of dementia compared to only 2% in the NSHAP pretest and NSHAP W2 samples. Their mean MoCA score was 19.8 (Cronbach alpha=0.858).

Item selection for the survey-adapted MoCA

Item selection was based on four criteria detailed in the methods section. Pretest data showed that the MoCA required 15.6 minutes (SD = 3.8) of survey administration time, requiring a reduction of at least 3.6 minutes to meet a goal of 12 minutes or less. Item difficulties and the percentage correct for each MoCA item in the pretest sample are reported in Supplementary Table 2. From the orientation domain, 4 items were excluded due to similarity to other orientation items or being too easy: day, year, place, and city. In the visuospatial skills domain, the cube item was excluded based on difficulty administering the item in the field and comparatively long administration time. From executive function, the abstraction item “train-bicycle” was removed due to similarity to the other abstraction item and it being the easier of the two. From the language domain, the camel and lion items were excluded due to being too easy, and the sentence item “John” was removed due to similarity to the other sentence item and difficulty administering in the field. Finally, from the attention domain, the vigilance item was excluded due to low item difficulty and difficult field administration. In summary, the following items were retained for the MoCA-SA: 1) Orientation: date and month (2 points total); 2) Executive function: abstraction - similarity of watch and ruler (1 point), modified trails-b (1 point); 3) Visuospatial skills: clock - contour, numbers, and hands (3 points total); 4) Memory: 5-word delayed recall (5 points); 5) Attention: forward digits (1 point), backward digits (1 point), subtract 7s (3 points); and 6) Language: naming rhinoceros (1 point), phonemic fluency - words with the letter “F” (1 point), and sentence repetition - “cat” (1 point). Total scores range from 0 to 20, with a mean administration time of 11.6 minutes in the pretest sample.

Relationship of MoCA-SA to original MoCA scores

We examined the relationship of the MoCA-SA to the original MoCA in the NSHAP pretest and clinic samples. The correlation between the survey-adapted and full scores was 0.97 in both samples. Regressing full MoCA on MoCA-SA scores in the NSHAP pretest yields the following equation to predict the MoCA score from the adapted score: $\text{MoCA} = 6.83 + (1.14 \times \text{MoCA-SA})$ (SE of intercept = 0.37; SE of slope = 0.02; RMSE = 0.994; Covariance of slope and intercept = -0.0089). The 95% prediction intervals were ± 0.10 to ± 0.25 points and 95% forecast intervals were approximately ± 2 points (Figure 1A).

We then estimated the slope of the prediction equation in the clinic sample to be 1.26 (SE=0.03), which was significantly different from the slope estimated in the pretest sample ($p=0.004$). After excluding individuals with dementia from the clinic sample ($n=21$), the estimated slope reduced slightly to 1.22 (SE=0.03), with a higher p -value ($p=0.08$). Thus, there is a small, but significant difference in the slopes of the prediction equations between the pretest and clinical samples, likely due in part to differences in the distribution of cognitive function between the two populations. Figure 1B shows the pretest prediction model applied to the clinical sample, with individuals with dementia shown as asterisks. The model performs well in individuals with no known diagnosis of dementia. When including individuals with dementia, the prediction model is less accurate at lower scores. The 18-item MoCA-SA has a Cronbach alpha of 0.740 in the NSHAP pretest sample, 0.806 in the clinical sample, and 0.773 in the full NSHAP W2 sample.

Unidimensional IRT (with partial credit) model

Item difficulties for the one-parameter and two-parameter partial credit models, together with estimated discrimination parameters for the latter, are shown in Table 1. A likelihood ratio test comparing the 2-parameter model to the 1-parameter model (assuming equal discrimination for all items) yields a p -value < 0.001 . Month (orientation), velvet and red (memory) and trails (executive function) had the highest discrimination, while discrimination was lowest for sentence and abstraction. Despite these differences, the correlation between the estimated cognitive abilities based on these two models was 0.99. Moreover, a simple sum of the MoCA-SA items yields a score very similar to the estimated abilities based on these models ($r = 0.99$ for the one-parameter partial credit model). Thus, a simple summed score of the MoCA-SA provides a good estimate of overall cognitive ability.

The distribution of cognitive abilities as estimated by the 1-parameter model are plotted on the log-odds scale (Figure 2, left). The distribution is slightly skewed toward lower abilities. The item difficulties, measured on the same scale, are also plotted (Figure 2, right) and their labels organized in columns according to cognitive domain. Item difficulties are distributed over the full range of abilities, with most located near the center of the distribution, as desired. Items from the visuospatial skills and attention domains cover the broadest range of the distribution, while those from the language, executive function and memory domains are primarily located in the middle. Orientation items are located near the bottom of the distribution, especially relevant for those with more severe impairment. Distinctions among the highest functioning individuals are captured primarily by the more difficult subtract 7s and memory items.

To determine whether items performed similarly when administered in different settings (survey vs. clinic) or by different interviewers (medical vs non-medical), as well as whether administering only the subset of 18 items changes item function, we compared item difficulties across the three samples. Of note, we do not exclude individuals with diagnoses of dementia for any of the three samples for the analysis. Overall, the profile of item difficulties was similar across the three samples (Figure 3). In all three samples, the temporal orientation (month, date) and clock contour items were the easiest, and the subtract 7s and delayed recall items were the most difficult, as judged either by the percent correct or the estimated item difficulties. However, there were some notable differences. The abstraction item was relatively more difficult for the W2 sample and less difficult for the NSHAP pretest. For the clinical sample, the subtract 7s item, clock hands, and digits forward, were relatively easier, while the naming (rhino) and trails-b items were non-significantly more difficult.

Item difficulties estimated from the W2 sample separately within selected demographic subgroups are plotted in Figure 4 (unlike the analyses above, these comparisons use a binary version of the subtract 7s item). In each case, the overall profiles are similar, though there were small, but statistically significant, differences for each ($p < 0.001$ in each case for a joint test of equal profiles). The subtract 7s items were relatively more difficult for women while the delayed recall items were more difficult for men. With the exception of small divergences in item difficulties for digits-forward and subtract 7s, the overall pattern of item

difficulties were similar in each racial/ethnic group. Subtract 7s and fluency items were relatively more difficult for those with less than a high school education.

DISCUSSION

We describe the development and performance of an 18-item survey-adaptation of the MoCA (MoCA-SA), which requires approximately 26% less administration time compared to the full MoCA. Items were retained from each cognitive domain assessed in the MoCA, and were selected to enhance reliable administration by non-medical interviewers while preserving discrimination across a broad range of cognitive abilities. MoCA-SA scores correlated highly with full MoCA scores in two independent samples, indicating that relatively little information is lost in utilizing the brief survey-adapted instrument to measure overall cognitive function. Consequently, a prediction model derived from our pretest sample allows accurate estimation of full MoCA scores in individuals with no known diagnosis of dementia.

In contrast to previous studies which have raised questions about the reliability of the MoCA in population-based samples,¹³ we found good internal reliability both for the full MoCA in our NSHAP pretest and for the 18-item MoCA-SA in the NSHAP W2 sample. There are at least two potential reasons for this apparent discrepancy. First, we likely preserved reliability through our efforts to minimize scoring in the field by non-medical interviewers and eliminate items that were most difficult to administer in this setting. Second, the greater variation in cognitive abilities among the US population of older, community-dwelling adults as compared to that in the populations studied by the previous reports may increase the proportion of true variation relative to error, yielding a higher reliability..

Additionally, in a sample of the general population of older adults, the 18 items from the MoCA-SA captured a broad range of cognitive abilities. For example, delayed-recall items and subtract 7s require relatively high levels of cognitive ability, orientation and visuospatial skills items require more limited cognitive function, and the remaining items have difficulties located throughout the center of the ability distribution. At the same time, eliminating several MoCA items and administering the remainder in a survey setting did not dramatically alter the individual items' functions. Specifically, the item difficulty profiles were similar when compared across NSHAP W2 (MoCA-SA, survey setting), NSHAP pretest (full MoCA, survey setting) and our clinical sample (full MoCA, administered by clinicians). One exception was the abstraction item, which was more difficult for those in the NSHAP W2 sample. This was most likely because the full MoCA contains a second, similar abstraction item administered *prior* to the one we retained, which provides an opportunity for respondents/subjects administered the full MoCA to become familiar with this type of item prior to being administered the second one.

Although most data in our study comes from administration of the shortened MoCA-SA in NSHAP W2, our results confirm and substantially extend what is known about the full MoCA's psychometric properties. Item difficulty profiles were similar across gender, major ethnic groups, and education, consistent with published analyses in clinical samples.¹⁵ Exceptions to this included the subtract 7s item (which was more difficult for women) and

the short-term memory items (which were more difficult for men), both of which are consistent with previous reports.^{26, 27} In addition, we found that those with less than a high school education had more difficulty with the subtract 7s and fluency items, perhaps because these items benefit from more mathematical or language training. Despite these minor differences, our results essentially confirm the effectiveness of the MoCA-SA (and by extension, the full MoCA) for capturing differences in cognitive functioning among a heterogeneous older population based on a large, probability sample of the U.S. population of older adults.

Developers of the MoCA selected items from several cognitive subdomains,¹² and previous psychometric work has shown a corresponding structure among the items.^{13, 15–17} However, its primary objective was to provide a clinical evaluation tool based on the overall score.¹² Likewise, our primary purpose in developing a survey-adapted version was to obtain an overall measure of cognitive function. Factor analyses (using a bi-factor model) on the Wave 2 sample (not reported here) are consistent with the presence of a generalized factor reflected in all of the items that accounts for the majority of the overall variation, and estimates of that factor are highly correlated with the simple summed score. However, further work with NSHAP and other datasets will be necessary to determine whether adequately reliable estimates of secondary factors can be obtained from the survey-form described here.

This study has limitations. First, our clinic sample is unique to the population served by the clinic, and the NSHAP pretest sample—while chosen to reflect a wide range of older adults based on age, sex and race/ethnicity—is not selected to be representative of any specific population. Thus, the results from those samples are less informative than those from the full NSHAP W2 sample. Second, the prediction model differed between the pretest and clinic samples when those in the clinic sample with dementia were included. As a result, the prediction equation should be used for populations comparable to the NSHAP population (i.e., community dwelling, and with sufficient cognitive function to complete the consent process and the interview). Further work will be needed to evaluate the prediction model in different samples and subgroups, particularly where dementia is highly prevalent. Third, the pretest and clinical sample sizes have limited power to compare item difficulties between these two samples. Thus, while this descriptive analysis provides a useful first step in comparing item performance between survey and clinical settings, additional work is required to confirm these findings. Nevertheless, the presence of differences between these settings would not invalidate the continued use of the full instrument in the clinic or of our short-form instrument in population-based samples. Fourth, we do not address the issue of clinical cutoffs, and how these might be applied to the U.S. population of community-dwelling adults and/or affected by survey administration and the shortened, survey-adapted instrument. There is considerable controversy regarding the appropriate cutoff for the full MoCA, particularly in community samples.²⁸ Because neuropsychological testing was not possible for the NSHAP samples and because many (if not most) of those with substantial cognitive impairment would not have met the criteria for participation (i.e., being able to provide informed consent and understand the instructions and questions), we were unable to address this issue. Further work in this area would help in comparing results from studies

using the MoCA in clinical settings to those using the MoCA (or the MoCA-SA) in population-based research.

In conclusion, our results demonstrate that a survey-adapted version of the MoCA, the MoCA-SA, can be successfully administered by non-medically trained interviewers as part of an omnibus, in-home health survey. The resulting scores are highly correlated with the full MoCA and retain the ability to discriminate across a broad range of cognitive abilities. Moreover, the measure functions similarly across major demographic subgroups in the U.S. Future research should be done to investigate the utility of establishing cutoffs for the MoCA-SA related to clinical categories, and to determine whether subscales can be extracted from the shortened instrument.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

source of funding:

This study was supported by the Clinical and Translational Science Award (CTSA) TL1 pre-doctoral training grant. This work was supported by funding for *MERIT* Award R37 AG030481 from the National Institute on Aging and the National Institutes of Health, including the National Institute on Aging, the Office of Women's Health Research, the Office of AIDS Research, the Office of Behavioral and Social Sciences Research, and the National Institute on Child Health and Human Development for the National Health, Social Life, and Aging Project (NSHAP R01AG021487, R37AG030481), and the NSHAP Wave 2 Partner Project (R01AG033903).

REFERENCES

1. Salthouse T. Consequences of age-related cognitive declines. *Annual Review of Psychology*. 2012; 63:201–226.
2. Pfeiffer E. A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatrics Society*. 1975; 23(10):433. [PubMed: 1159263]
3. Tombaugh TN, McIntyre NJ. The mini-mental state examination: a comprehensive review. *Journal of the American Geriatrics Society*. 1992; 40(9):922–935. [PubMed: 1512391]
4. Friedman TW, Yelland GW, Robinson SR. Subtle cognitive impairment in elders with Mini-Mental State Examination scores within the 'normal' range. *International journal of geriatric psychiatry*. 2011; 27(5):463–471. [PubMed: 21626569]
5. Pendlebury ST, Cuthbertson FC, Welch SJ, Mehta Z, Rothwell PM. Underestimation of Cognitive Impairment by Mini-Mental State Examination Versus the Montreal Cognitive Assessment in Patients With Transient Ischemic Attack and Stroke A Population-Based Study. *Stroke*. 2010; 41(6):1290–1293. [PubMed: 20378863]
6. Shega J, Sunkara P, Kotwal A, et al. Measuring Cognition: The Chicago Cognitive Function Measure (CCFM) in the National Social Life, Health and Aging Project (NSHAP), Wave 2. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2013 IN PRESS.
7. Brandt J, Spencer M, Folstein M. The telephone interview for cognitive status. *Cognitive and Behavioral Neurology*. 1988; 1(2):111–118.
8. Pressley JC, Trott C, Tang M, Durkin M, Stern Y. Dementia in community-dwelling elderly patients: a comparison of survey data, medicare claims, cognitive screening, reported symptoms, and activity limitations. *Journal of clinical epidemiology*. 2003; 56(9):896–905. [PubMed: 14505776]
9. Shulman KI. Clock-drawing: is it the ideal cognitive screening test? *International journal of geriatric psychiatry*. 2000; 15(6):548–561. [PubMed: 10861923]

10. Johnson JK, Lui L-Y, Yaffe K. Executive function, more than global cognition, predicts functional decline and mortality in elderly women. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 2007; 62(10):1134–1141.
11. Blacker D, Lee H, Muzikansky A, et al. Neuropsychological measures in normal individuals that predict subsequent cognitive decline. *Archives of Neurology*. 2007; 64(6):862. [PubMed: 17562935]
12. Nasreddine ZS, Phillips NA, Bédirian V, et al. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*. 2005; 53(4):695–699. [PubMed: 15817019]
13. Bernstein IH, Lacritz L, Barlow CE, Weiner MF, DeFina LF. Psychometric evaluation of the montreal cognitive assessment (MoCA) in three diverse samples. *The Clinical Neuropsychologist*. 2011; 25(1):119–126. [PubMed: 21154110]
14. Subar AF, Ziegler RG, Thompson FE, et al. Is shorter always better? Relative importance of questionnaire length and cognitive ease on response rates and data quality for two dietary questionnaires. *American journal of epidemiology*. 2001; 153(4):404–409. [PubMed: 11207159]
15. Koski L, Xie H, Finch L. Measuring cognition in a geriatric outpatient clinic: Rasch analysis of the Montreal Cognitive Assessment. *Journal of geriatric psychiatry and neurology*. 2009; 22(3):151–160. [PubMed: 19307324]
16. Freitas S, Simoes MR, Marôco J, Alves L, Santana I. Construct validity of the Montreal Cognitive Assessment (MoCA). *J Int Neuropsychol Soc*. 2012; 18(2):1–9. [PubMed: 22152338]
17. Duro D, Simões MR, Ponciano E, Santana I. Validation studies of the Portuguese experimental version of the Montreal Cognitive Assessment (MoCA): confirmatory factor analysis. *Journal of neurology*. 2010; 257(5):728–734. [PubMed: 19937448]
18. Tsai C-F, Lee W-J, Wang S-J, Shia B-C, Nasreddine Z, Fuh J-L. Psychometrics of the Montreal Cognitive Assessment (MoCA) and its subscales: validation of the Taiwanese version of the MoCA and an item response theory analysis. *International Psychogeriatrics*. 2012; 24(04):651–658. [PubMed: 22152127]
19. NORC. National Social Life, Health, and Aging Project (NSHAP). 2013 <http://www.norc.org/Research/Projects/Pages/national-social-life-health-and-aging-project.aspx>.
20. Smith S, Jaszczak A, Graber J, et al. Instrument development, study design implementation, and survey conduct for the National Social Life, Health, and Aging Project. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2009; 64(suppl 1):i20.
21. Weisberg, S. *Applied linear regression*. Vol. 528. Wiley; 2005.
22. Rasch, G. *Probabilistic Models for Some Intelligence and Attainment Tests*. Vol. 1. University of Chicago Press; 1980.
23. Hays RD, Morales LS, Reise SP. Item response theory and health outcomes measurement in the 21st century. *Medical Care*. 2000; 38(9 Suppl):II28. [PubMed: 10982088]
24. Rabe-Hesketh S, Skrondal A, Pickles A. GLLAMM manual. 2004 <http://biostats.bepress.com/ucbbiostat/paper160/>.
25. Stata Statistical Software: Release 12 [computer program]. College Station, TX: StataCorp LP; 2011.
26. Rosselli M, Tappen R, Williams C, Salvatierra J. The relation of education and gender on the attention items of the Mini-Mental State Examination in Spanish speaking Hispanic elders. *Archives of clinical neuropsychology*. 2006; 21(7):677–686. [PubMed: 16968662]
27. Jones RN, Gallo JJ. Education and sex differences in the mini-mental state examination effects of differential item functioning. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*. 2002; 57(6):P548–P558.
28. Rossetti HC, Lacritz LH, Cullum CM, Weiner MF. Normative data for the Montreal Cognitive Assessment (MoCA) in a population-based sample. *Neurology*. 2011; 77(13):1272–1275. [PubMed: 21917776]

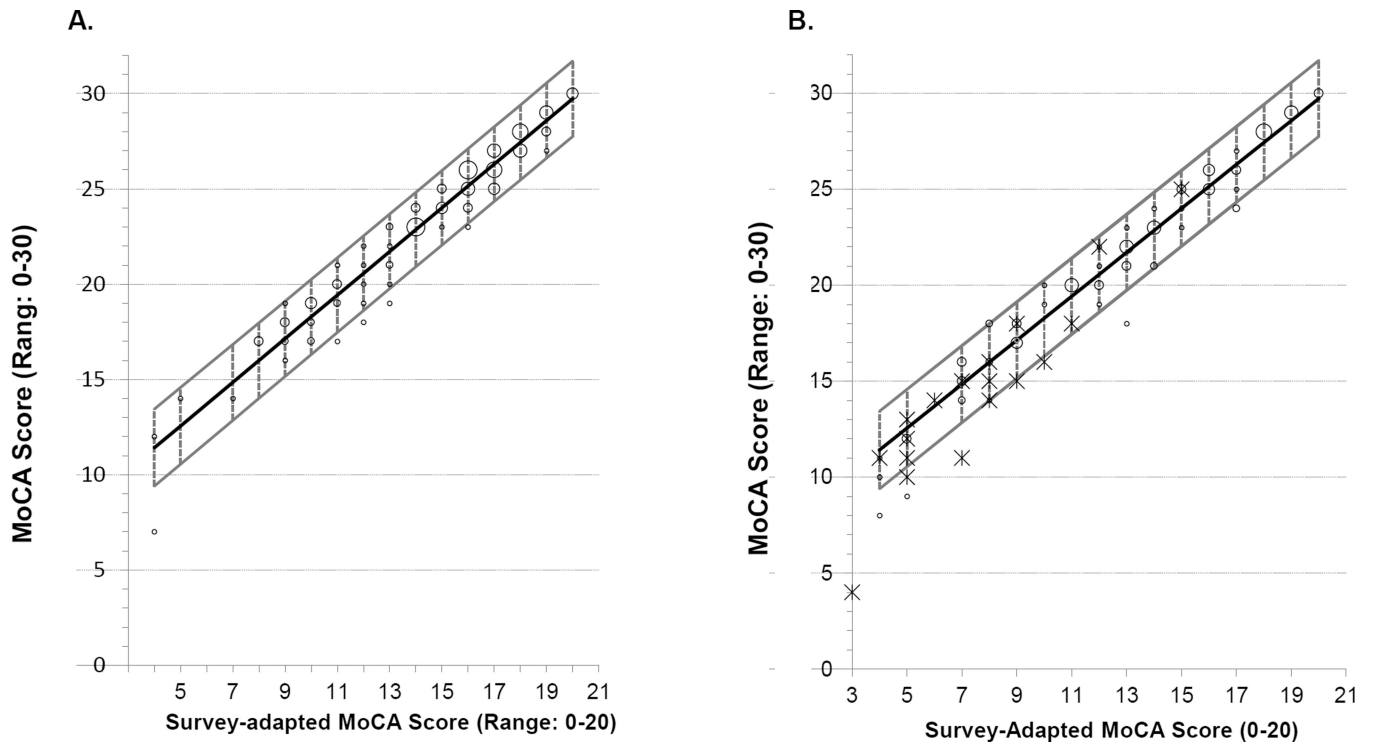


Figure 1. Relationship of 18-item survey-adapted MoCA to the 28-item MoCA. A) Prediction Model with 95% Forecast Interval in NSHAP Pretest (n=120). B) Prediction Model in an independent clinical sample (n=93). Abbreviations: MoCA – Montreal Cognitive Assessment. Larger points represent more individuals with that observation. Individuals with dementia are shown as asterisks.

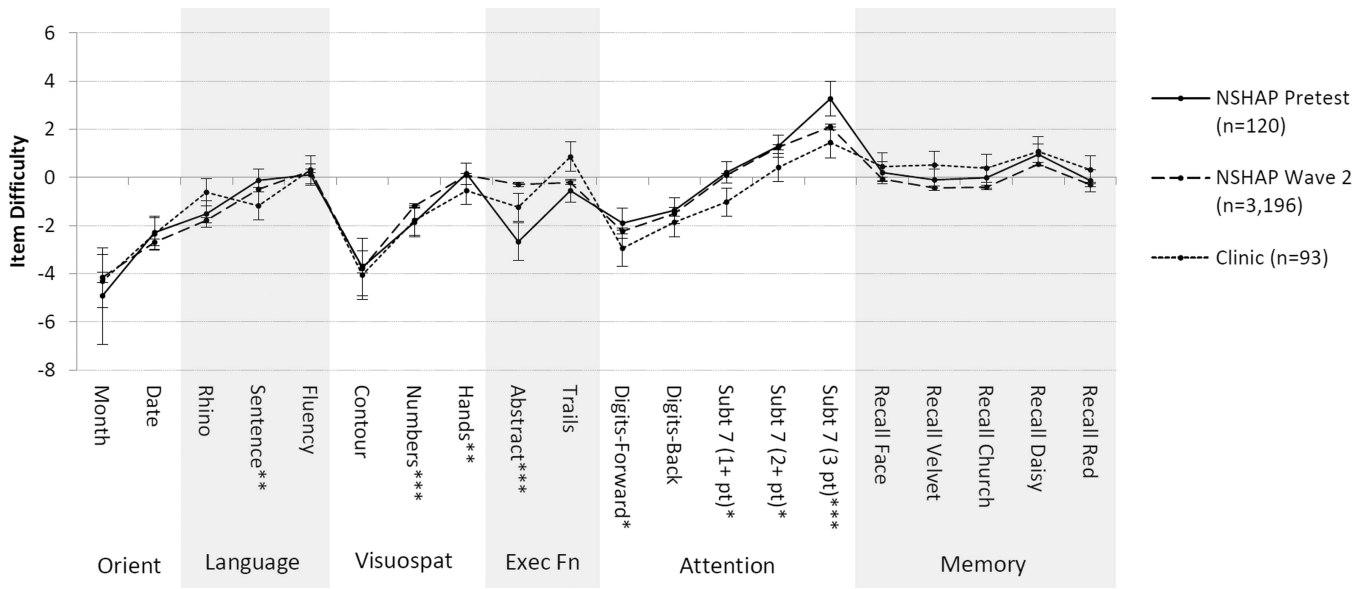


Figure 3. Comparison of Item Difficulties in NSHAP Pretest, NSHAP Wave 2, and Clinic Samples. Abbreviations: NSHAP - National Social life Health and Aging Project, Subt 7 – Subtract 7s, Orient – Orientation, Visuospat – Visuospatial skills, Exec Fn – Executive Function. P-values from joint Wald tests are represented by *** for $p < 0.001$, ** for $p < 0.01$, and * for $p < 0.05$. Error bars represent 95% confidence intervals.

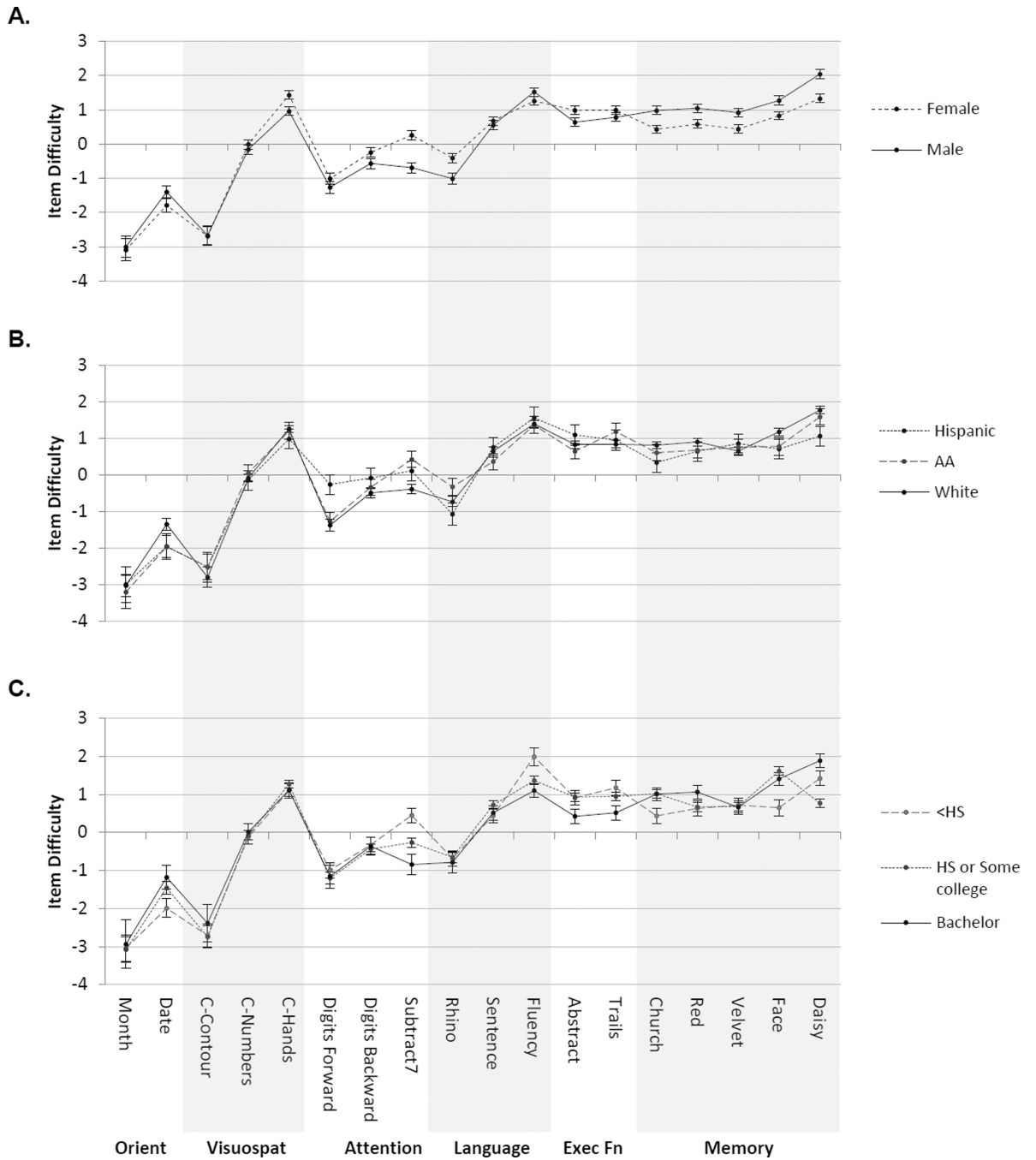


Figure 4. Comparison of Item Difficulties in NSHAP Wave 2 Demographic Subgroups Including A) Gender, B) Ethnicity, and C) Education. Abbreviations: NSHAP - National Social life Health and Aging Project, C - Clock, Orient – Orientation, Visuospat – Visuospatial skills, Exec Fn – Executive Function. Error bars represent 95% Confidence Intervals.

Table 1

Item difficulties for the MoCA-SA in NSHAP Wave 2 as estimated by the 1-parameter and 2-parameter Partial-Credit Model (n=3,196)

| Domain | Item | % Correct ^a | 1-parameter model | 2-parameter model | |
|--|-----------------|------------------------|------------------------------|------------------------------|--------------------------|
| | | | Difficulty (SE) ^b | Difficulty (SE) ^b | Item discrimination (SE) |
| Temporal Orientation | Month | 97.2 | -4.16 (0.11) | -5.65 (0.34) | 1.00 |
| | Date | 90.2 | -2.69 (0.07) | -2.67 (0.09) | 0.49 (0.05) |
| Language | Rhino | 81.1 | -1.79 (0.05) | -1.75 (0.06) | 0.47 (0.05) |
| | Sentence | 59.9 | -0.49 (0.05) | -0.45 (0.04) | 0.34 (0.04) |
| | Fluency | 44.8 | 0.27 (0.04) | 0.29 (0.05) | 0.58 (0.06) |
| Visuospatial Skills (Clock Draw) | Contour | 96.1 | -3.79 (0.10) | -3.69 (0.14) | 0.46 (0.06) |
| | Numbers | 72.3 | -1.19 (0.05) | -1.18 (0.05) | 0.49 (0.05) |
| | Hands | 48.4 | 0.09 (0.04) | 0.08 (0.04) | 0.44 (0.05) |
| Executive Function | Abstraction? | 56.0 | -0.29 (0.04) | -0.27 (0.04) | 0.33 (0.04) |
| | Trails | 54.6 | -0.22 (0.04) | -0.24 (0.05) | 0.65 (0.07) |
| Attention, Concentration, and Working Memory | Digits-Forward | 86.1 | -2.24 (0.06) | -2.16 (0.07) | 0.45 (0.05) |
| | Digits-Backward | 77.2 | -1.51 (0.05) | -1.44 (0.05) | 0.44 (0.05) |
| | Subtract 7s | | | | |
| | 1+ point | 84.7 | 0.09 (0.04) | 0.10 (0.05) | 0.57 (0.06) |
| | 2+ points | 73.7 | 1.25 (0.05) | 1.30 (0.05) | - |
| | 3 points | 52.5 | 2.10 (0.06) | 2.18 (0.07) | - |
| Short-term Memory (Delayed Recall) | Face | 51.6 | -0.07 (0.04) | -0.07 (0.04) | 0.48 (0.05) |
| | Velvet | 58.9 | -0.44 (0.05) | -0.48 (0.05) | 0.67 (0.07) |
| | Church | 58.3 | -0.41 (0.05) | -0.41 (0.05) | 0.53 (0.06) |
| | Daisy | 39.6 | 0.54 (0.05) | 0.56 (0.05) | 0.56 (0.06) |
| | Red | 56.2 | -0.30 (0.04) | -0.32 (0.05) | 0.63 (0.07) |

Abbreviations: NSHAP - National Social life Health and Aging Project (NSHAP), SE - Standard error, MoCA-SA – Survey-Adapted Montreal Cognitive Assessment

^a % Correct represents percentage of sample obtaining correct response

^b Item difficulty represents the level of cognitive ability needed to have a 50% probability of responding to the item correctly