Measuring Cognition: The Chicago Cognitive Function Measure in the National Social Life, Health and Aging Project, Wave 2

Joseph W. Shega,[1] Priya D. Sunkara,[2] Ashwin Kotwal,[3] David W. Kern,[4] Sara L. Henning,[5] Martha K. McClintock,[6] Philip Schumm,[3] Linda J. Waite,[7] and William Dale[2]

[1]VITAS Innovative Hospice Care, Maitland, Florida.
[2]Department of Medicine,
[3]Department of Health Studies,
[4]Department of Comparative Human Development,
[5]NORC at the University of Chicago,
[6]Department of Psychology, and
[7]Department of Sociology, University of Chicago, Illinois.

Objectives. To describe the development of a multidimensional test of cognition for the National Social life, Health and Aging Project (NSHAP), the Chicago Cognitive Function Measure (CCFM).

Method. CCFM development included 3 steps: (a) A pilot test of the Montreal Cognitive Assessment (MoCA) to create a standard protocol, choose specific items, reorder items, and improve clarity; (b) integration into a CAPI-based format; and (c) evaluation of the performance of the CCFM in the field. The CCFM was subsequently incorporated into NSHAP, Wave 2 ($n = 3,377$).

Results. The pre-test ($n = 120$) mean age was 71.35 (SD 8.40); 53% were female, 69% white, and 70% with college or greater education. The MoCA took an average of 15.6 min; the time for the CCFM was 12.0 min. CCFM scores (0–20) can be used as a continuous outcome or to adjust for cognition in a multivariable analysis. CCFM scores were highly correlated with MoCA scores ($r = .973$). Modeling projects MoCA scores from CCFM scores using the equation: MoCA = (1.14 × CCFM) + 6.83. In Wave 2, the overall weighted mean CCFM score was 13.9 (SE 0.13).

Discussion. A survey-based adaptation of the MoCA was successfully integrated into a nationally representative sample of older adults, NSHAP Wave 2.

Key Words: Cognitive assessment—Health—Older adults—Social factors.

COGNITION represents a key component of health, and decrements in cognitive functioning commonly accompany advancing age (Salthouse, 2012; Steinerman, Hall, Sliwinski, & Lipton, 2010). The individual cognitive domains affected (e.g., memory, attention, or executive function) vary within an individual and across a population by age (Salthouse, 2012). The subtler variations in cognitive functioning among community-dwelling older adults may represent relevant and important determinants of overall health status Moreover, the modest changes in cognition that do occur, if they progress, may advance to mild cognitive impairment (MCI; deficits in at least one cognitive domain, characteristically memory, but without clear functional consequences) and/or overt dementia (deficits in at least two cognitive domains of which one is memory with functional consequences) (Ashford, 2008; Petersen et al., 1999).

A paucity of evidence exists surrounding the relationship between "normal" cognitive functioning and health in older adults (Alwin & Hofer, 2011). Emergent research suggests variation in cognitive performance in individuals with no known (or subclinical) cognitive impairment is associated with psychological, physical, and social health. For example, one study found that cognitively intact older adults who had lower scores on a measure of global cognitive function reported more depressive symptoms, whereas those with higher cognitive scores reported fewer such symptoms (Santos et al., 2013). Similarly, lower cognitive performance has been associated with physical disability and poorer social engagement compared to persons with better cognitive performance among cognitively intact individuals (Ishizaki et al., 2006; Paulo et al., 2011). Finally, initial studies investigating the relationship between cognitive performance with sleep quality and duration are on-going, with the former appearing more important to function than the latter (Blackwell et al., 2006; Potvin et al., 2012). While most of these studies are cross-sectional, limiting the ability to infer causality, and do not account for treatment effects (e.g., medication use), available studies highlight the importance of integrating multidimensional cognitive measures into population-based epidemiologic research.

Epidemiologic studies traditionally incorporate some measure of cognition, although study design (e.g., phone vs in-person interview) or competing priorities within the survey may limit the number and extent of cognitive domains assessed (Herzog & Wallace, 1997). For example, the Health and Retirement Study (HRS) evaluates verbal memory, orientation, numeracy, attention, reasoning ability, vocabulary, verbal fluency, and language, but overlooks visuo-construction and executive function (Wallace & Herzog, 1995). As part of The National Social life, Health and Aging Project (NSHAP)—a population-based, nationally representative, in-home study of community-dwelling older adults—we developed a more comprehensive cognitive function measure as part of Wave 2. After an extensive literature review of available cognitive measures, the Montreal Cognitive Assessment (MoCA) stood out as a promising starting point to build upon in creating a measure, because it assesses key cognitive domains (executive function, visuo-construction skills, naming, memory, attention, language, abstract thinking, and orientation), has been administered in a variety of settings, and does not take an inordinate amount of time to complete (Luis, Keegan, & Mullan, 2009; Nasreddine et al., 2005).

Our purpose here is to describe an adaptation of the clinical MoCA for use in a survey setting, the NSHAP, Wave 2, and provide recommendations for its use. This includes pilot testing to revise the instructions and format of the measure to ensure it could be reliably administered by non-medically trained personnel (e.g., field interviewers) in a home setting. Respondent burden was also minimized through elimination of redundant items, the use of an enhanced visual interface to better engage participants, and modification of the item order. These changes helped encourage completion of the measure as well as the remainder of the full 2-hr NSHAP interview. Using a pre-test sample, the measure was further refined for incorporation into Computer-Assisted Personal Interviewing (CAPI) technology for the full survey. With consideration of several guiding principles, outlined below, a final adaption, identified as the Chicago Cognitive Function Measure (CCFM), was constructed and incorporated into the NSHAP, Wave 2 instrument. In addition, a scoring algorithm for the measure was created and applied, yielding the final measure accompanied by recommendations on its use. Finally, the CCFM was successfully incorporated into NSHAP Wave 2 with preliminary results being reported.

Method

Overview

NSHAP, conducted in conjunction with NORC at the university of Chicago (http://www.norc.org), is a population-based, nationally representative, in-home longitudinal survey study designed to investigate the complex interplay of social, biological, emotional, and environmental factors

that come with aging. The data collection process included three components: (a) an in-home, face-to-face interview; (b) a collection of biomeasures; and (c) a leave-behind, self-administered questionnaire. The first wave of NSHAP, conducted in 2005–2006, surveyed older adults aged 57–85, and only included participants who could provide informed consent. In Wave 1, cognition was assessed in the face-to-face interview using the Short Portable Mental Status Questionnaire (SPMSQ), a 10-question cognitive screening measure originally designed to identify "organic brain deficits" (Pfeiffer, 1975). Applying a cumulative error approach from the literature, 96% of Wave 1 participants were designated as having "normal" cognitive functioning (Special Issue 1). Data are publicly available (NSHAP Wave 1: Waite, Linda J., Edward O. Laumann, Wendy Levinson, Stacy Tessler Lindau, and Colm A. O'Muircheartaigh. National Social Life, Health, and Aging Project (NSHAP): Wave 1. ICPSR20541-v6. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-04-30. doi:10.3886/ICPSR20541.v6.).

Given the apparent "ceiling effect" of the SPMSQ for a community-dwelling older population, the Wave 2 investigative team sought to incorporate a cognitive measure that captured overall global functioning and had intra- and inter-individual variability among cognitively intact individuals across multiple domains (Rossetti, Lacritz, Cullum, & Weiner, 2011). Following a literature review, the MoCA was chosen as a starting framework for the CCFM, because it demonstrated reliability and validity in several clinical settings, takes about 15 min to complete, and assesses cognitive domains of interest to NSHAP investigators, including those not traditionally incorporated into population-based studies such as visuo-construction skills and executive function. The cognitive domains, with individual items in parentheses, are presented to respondents in the following order: executive function (Trails b), visuo-construction skills (cube and clock), naming (animal recognition), memory immediate recall (five words), attention (forward digit span, backward digit span, vigilance task, and serial 7s), language (sentence repetition and verbal fluency), abstraction (similarity between items), memory delayed recall (five words), and orientation. Composite MoCA scores (0–30 points) sums domain performance and reflects global cognitive functioning. Higher scores indicate better cognition. Tabulated scores within individual domains can be used to describe relative cognitive strengths and weaknesses.

The development of the NSHAP CCFM included three major steps: (a) Administration of a pilot test of the MoCA items with cognitive interviewing to refine a standard protocol, choose the items to be administered, reorder items, and adjust the instructions for improved clarity; (b) integration of the selected questions into a CAPI-based format that could be used in the home; and (c) a pre-test (*n* = 120) to evaluate the performance of the resulting measure in the field, with development of a formal scoring protocol. The pre-test also

provided key information to evaluate each item and determine which should be incorporated into the CCFM. The CCFM was subsequently incorporated into NSHAP Wave 2 (*n* = 3,377). The pilot, pre-test, and NSHAP Wave 2 were approved by the University of Chicago Institutional Review Board.

Pilot Test

Participants were recruited from the University of Chicago's South Shore Senior Center, a community-based, demographically diverse, geriatrics clinic on Chicago's south side. All participants were age 65 or older and designated as cognitively intact by their primary care provider. We administered revised versions of the cognitive measure until no new themes emerged surrounding our main objectives. These objectives included: (a) To optimize question wording to facilitate administration of the measure by NSHAP's field interviewers in a participant's home; (b) To evaluate item ordering to maximize completion rates; (c) To modify the layout from a one-page format to one that could be administered with CAPI. A trained researcher with an interest in cognitive aging applied cognitive interviewing techniques and direct observation of participant reaction in order to address each of the listed objectives.

Respondents were read the instructions for each item and asked to evaluate for comprehension (Nasreddine, 2003). Many participants reported difficulty understanding the directions to the trails-b task. The instructions in the original MoCA manual read, "Please draw a line, going from a number to a letter in ascending order…" A simple substitution of the word "ascending" with "increasing" led to improved comprehension of the instructions and higher completion rates for the item. When probed about ways to facilitate administering the items, participants responded favorably to the incorporation of "sign posts" to signal when the cognitive assessment was beginning and that the items would vary by level of difficulty. Instructions were also modified to better account for the measure being administered in a participant's home. For example, to reduce the frequency of respondents searching for answers in their immediate surroundings (e.g., clocks or calendars), field interviewers were instructed to say, "Try your best without using clues from around the room." Also, we wanted to minimize the contribution of sensory deficits to the results, so the instructions included items such as, "wear your glasses if needed for reading."

There is no known "order effects" among individual MoCA items (Nazem et al., 2009). However, cognitive interviewing demonstrated that item placement directly influenced respondent likelihood of further participation. The usual presentation of the challenging "trails-b" as the first task intimidated many participants, who perceived it as overwhelming and difficult to follow. Participants reported being most comfortable attempting the items when the relatively easier "orientation questions" (e.g., day, time) were administered first, which led to a smoother transition for the remainder of the tasks. Orientation was followed by animal

naming, and the visuo-spatial domain (e.g., clock draw, copy cube, and trails-b). The remainder of the items was administered in the order dictated by the MoCA instruction booklet in order to best maintain the time interval between immediate and delayed recall.

The final objective of the pilot test was to modify the traditional one-page pen and paper MoCA layout to a version conducive to an in-home setting, with a survey instrument using CAPI. Cognitive interviewing found respondents were more likely to complete many of the items (e.g., pictures) when they were enlarged and presented individually on separate index cards. Respondents reported the smaller images to be intimidating, while larger images put respondents at ease and mitigated mistakes in item completion secondary to poor vision. For example, participants performed better on the animal naming tasks when larger images were incorporated. Also, respondents reported it being much easier to write and discern the numbers written on the clock face when more space was provided. Moreover, based upon NSHAP Wave 1 experience, enlarged items would facilitate test administration in the field when interviews, on occasion, are in environments that requires a larger-than-typical distance between the interviewer and respondent (e.g., having no table available to conduct the interview, which led the interviewer and respondent to sit on opposite ends of a couch).

In clinical settings, the MoCA is scored by the clinician administering the test. While conducting the pilot testing, concern arose surrounding the reliability of the field interviewers to administer items while simultaneously recording responses and scoring items accurately, something demonstrated in previous studies (Price et al., 2011). For example, interviewers could disagree on whether or not each of the three points allocated to the clock item (i.e., contour, numbers, and hand) were correct. Also, the responsibility of scoring necessarily detracts from documentation of the task, something crucial for field interviewing. For instance, the verbal fluency task queries respondents to name as many words as they can think of that begin with the letter "F" in 60 s. The responsibility of scoring these rapidly stated answer can shift the interviewer's focus strictly to the number of words stated, and consequently to overlook the scoring guidelines disallowing proper nouns, numbers, or words that begin with the same root, but a different suffix (e.g., live, lived, living). As a result, efforts were focused on reliable documentation of the cognitive measures in the field, and scoring was conducted post-data collection by different personnel rigorously trained on item scoring. For instance, field interviewers recorded, on a pre-numbered sheet, all of the words produced by respondents within the 60 s time frame, response regardless of "correctness." This ensured that as much data as possible was returned to the research team.

Adaptation and Integration into CAPI

The content and structure of the survey questions established in the pilot test were programmed into CAPI. Successful integration of the CCFM into CAPI required that the measure

assimilate into the rest of the NSHAP survey, that the CAPI communicate effective and standard instructions, and that it collect meaningful data. The CAPI provided a detailed script and prompts which, in conjunction with the protocol booklet that was part of the field interviewers' training manual, ensured uniformity across interviewers. Unique programming features were created to reinforce this uniformity and to decrease interviewer burden. For example, in the orientation questions, the CAPI prompted the interviewer to ask, "What is the day today?" and for subsequent orientation questions, CAPI auto-filled the correct date, so that when interviewers entered the participant's response, it could be readily checked for correctness. The CAPI could then be used to check whether the item was correct or incorrect based on the information stored for that particular interview and time stamps for the date.

A number of other items were aided by the creation of sufficient response categories in CAPI. For example, the abstraction task included common incorrect answers as reported in the original MoCA literature as well as common incorrect answers that arose in the pilot test. The pre-test also provided additional incorrect answers to be included in the final Wave 2 survey.

Programming features were also implemented to account for more difficult tasks, such as the "serial 7s" task. Interviewers were instructed to enter six numbers into CAPI, instead of the five on the clinical version, to eliminate possible data entry mistakes in cases where the respondent first answered "100"—which was not counted as a correct answer. That way, scorers could later evaluate each response individually. In a hypothetical sequence, where the first response given is "90," which is incorrect, and the next number is "83," the "83" would be marked as a correct response even though it was not generated by subtracting multiples of 7 from 100.

Items on the pretest that were captured by pencil and paper tasks included animal recognition, cube and clock, trails b, vigilance task (letter sequence), and verbal fluency. The remainder of the items (memory immediate and delayed recall, forward digit span, backward digit span, serial 7s, sentence repetition, similarity between items, and orientation were administered and documented via CAPI.

Pre-Test

Prior to being fielded, a pre-test version of the NSHAP Wave 2 survey was administered. The pre-test sample was a combination of a list sample provided to NORC by a third party vendor along with retests of the NSHAP Wave 1 pre-test participants, yielding 120 cases altogether.

The pre-test included all of the MoCA questions in their revised form based on the pilot test, embedded within the entire proposed NSHAP in-home interview in CAPI format. Additional goals of the pre-test were to ensure the proper administration of the items by the field interviewers, as well as to document any issues that arose during the administration of the MoCA. A text box was provided at the end of the

MoCA in CAPI for field interviewers to make additional notes. Also, the pre-test data was used for item selection for the final creation of the CCFM.

Evaluation CCFM in the Field, NSHAP Wave 2

NSHAP is a nationally representative probability sample of community-dwelling older adults without known cognitive impairment. The CCFM was available for use in English and Spanish among NSHAP Wave 2 participants. Wave 2 individuals were enrolled in 2010–2011, with a weighted response rate of 76.9% ($n = 3,377$). Data are publicly available (NSHAP Wave 2: Waite, Linda J., Kathleen Cagney, William Dale, Elbert Huang, Edward O. Laumann, Martha K. McClintock, Colm A. O'Muircheartaigh, L. Phillip Schumm, and Benjamin Cornwell. National Social Life, Health, and Aging Project (NSHAP): Wave 2 and Partner Data Collection. ICPSR34921-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2014-04-29. doi:10.3886/ICPSR34921.v1.).

Data Analysis

Descriptive statistics were summarized as means and standard deviations for continuous variables, and percentages for categorical variables. A new time-stamp was automatically created in CAPI every time a new task was started. The time to complete each item was recorded (via time stamps) as the average time in minutes; this included the time to administer the instructions, to get participant responses, and to document those responses. The orientation section was divided into the following three groups (a) "month," (b) "date," "year," and "day," and (c) "place" and "city." The proportion of participants who attempted to complete each CFM item, those who completed it, and those who completed it correctly was tabulated. Total scores were tabulated using published guidelines without adjustment for education given that this was a non-clinical research sample and education adjustments continue to be debated (Bernstein, Lacritz, Barlow, Weiner, & DeFina, 2011; Nasreddine, 2003; Nasreddine et al., 2005) The research group subsequently examined pre-test data to determine which items to include in the NSHAP-Wave 2 survey as the CCFM. This was done in conjunction with guiding principles for selection developed by the research team a priori. These principles include: (a) question representation from each of the eight cognitive domains of interest, (b) inclusion of items that were difficult enough to ensure variability of response, yet not so difficult to impact interviewer rapport or deter completion of the remainder of the interview, (c) the ability to be administered via CAPI by field interviewers, and (d) administration time that did not exceed 12 min. CCFM performance for NSHAP Wave 2 is reported as "percent correct" separated by domain and item and weighted by sampling strategy. Weighted mean and standard errors for CCFM performance is also reported. Linear regression modeling was utilized to project MoCA scores from CCFM scores from NSHAP Wave 2 based upon the pre-test data.

RESULTS

A total of 120 interviews were conducted in participant homes as part of the pre-test. This included 25 (21%) participants from NSHAP Wave 1, 26 (22%) interviews with respondent partners, and 69 (57%) from List Sample respondents provided by NORC. Participant demographics are displayed in Table 1.

Table 1. Pre-Test Participant Demographics, *N* = 120

Characteristic	Value
Age, mean (*SD*)	71.35 (8.40)
Female, %	52.5
Ethnicity, %	
White	68.91
Black	12.61
Hispanic, non-black	14.29
Other	4.20
Education, %	
Less than high school	7.50
High school	22.50
Any college	35.00
College degree or more	35.00
Household income, %	
$25,000 or less	24.78
$25,001–$50,000	38.94
$50,001–$100,000	27.43
$100,001 or greater	8.85
Currently working, %	27.73
Co-morbid conditions, %	
Arthritis	30.00
Hypertension	67.50
Heart condition	25.83
Cancer (other than skin)	16.67
Diabetes	30.00
Emphysema or asthma	20.00
Stroke	9.17
Parkinson's	1.67
Osteoporosis	13.45

Each participant completed all items including CAPI generated assessments and "pencil and paper" items. CAPI was successful in delivering time data for all 120 interviews; results are displayed in Table 2 by order of administration. The times listed for each item include the time to administer the instructions, to record the participant response, and to document the results by the interviewer. Overall, the full MoCA took an average of 15.63 min to complete, ranging from 8.92 to 26.93 min, with a standard deviation of 3.80 min. In general, the orientation items took the least amount of time to complete, whereas fluency took the longest (2.23 min). The process of encoding five words for delayed recall had the greatest range of response, 0.72–5.47 min (1.58 min average).

Table 3 reports individual survey items, displaying the proportion of participants who attempted each item, the proportion that completed it, and the proportion completing it correctly. Few participants refused to attempt individual items, with the highest refusal rate being 3.3%. The completion rate for individual items ranged from 74.0%–100.0%, with delayed recall having the lowest completion rate. The proportion of respondents who correctly completed individual items varied from 40.0% to 99.2%. Within each domain by proportion of who completed it correctly, the items had the following performance: (a) orientation—"year" 90.0% correct; (b) animal naming—"rhinoceros" 83.3% correct; (c) executive—"trails b" 69.17% correct; (d) visuo-construction—"cube" 47.5% correct; (e) attention—"serial 7s" 61.7% correct; (f) language—"fluency" 56.7% correct; (g) abstraction—"similar train/bike" 65.8% correct; and (h) delayed recall—"daisy" 40.0% correct. Total MoCA scores ranged from 7 to 30 as displayed in Table 4.

The items to include in the CCFM were subsequently selected based upon our guiding principles (described in Method section). For example, the decision to include the

Table 2. Pre-Test MoCA Administration Time by Item in Minutes, Including the Time to Administer the Instructions, Participant Response, and Its Documentation by the Interviewer, *N* = 120

Domain	MoCA item	Mean time (min)	Range (min)	Standard deviation (min)
Orientation	Month/date/year	1.25	0.53–3.35	0.57
	Day of week	0.15	0.07–0.48	0.06
	Place/city	0.49	0.27–1.12	0.17
Naming	Animal naming	0.64	0.22–2.58	0.37
Executive function	Trails-b	1.43	0.12–6.57	0.89
Visuo-construction	Clock	1.22	0.30–3.58	0.58
	Cube	0.97	0.07–2.67	0.60
Memory	Encode five words	1.58	0.72–5.47	0.75
	Delayed recall	0.62	0.23–1.83	0.27
Attention	Forward digits	0.34	0.20–1.10	0.12
	Backward digits	0.31	0.17–0.75	0.10
	Vigilance (As)	1.19	0.75–3.85	0.36
	Serial 7s	1.54	0.57–4.13	0.66
Language	Sentence repeat (cat)	0.43	0.17–1.83	0.24
	Sentence repeat (John)	0.43	0.17–1.83	0.24
	Fluency (F's)	2.23	1.22–4.63	0.52
Abstraction	Similarity (train/bicycle)	0.19	0.07–0.65	0.11
	Similarity (watch/ruler)	0.19	0.07–0.95	0.13

Note. MoCA = Montreal Cognitive Assessment.

Table 3. Pre-Test MoCA Items Attempted, Completed, and Completed Correctly, *N* = 120

Domain	Item (points ascribed)	Attempted response, %	Complete response, %	Response correct, %
Orientation	Month (1)	100	100	99.17
	Date (1)	98.33	98.33	99.17
	Year (1)	99.17	99.17	90.00
	Day (1)	100	100	98.33
	Place (1)	99.17	99.17	94.17
	City (1)	100	100	96.67
Naming	Lion (1)	99.17	99.17	98.33
	Rhinoceros (1)	99.17	99.17	83.33
	Camel (1)	99.17	99.17	97.50
Visuo-construction	Clock contour (1)	100	100	97.48
	Clock numbers (1)	100	100	87.30
	Clock hands (1)	100	92.50	56.30
	Copy cube (1)	100	89.17	47.50
Executive function	Trails B (1)	96.67	86.67	69.17
Memory: immediate five-word recall	Face (0)	100	[a]	[a]
	Velvet (0)	100		
	Church (0)	100		
	Daisy (0)	100		
	Red (0)	100		
Memory: delayed five-word recall	Face (1)	100	74.00[b]	55.00
	Velvet (1)	100		60.83
	Church (1)	100		59.17
	Daisy (1)	100		40.00
	Red (1)	100		61.67
Attention	Forward digits (1)	100	100	87.50
	Backward digits (1)	100	100	81.67
	Vigilance "A" (1)	100	98.33	90.83
	Serial 7s			
	0 correct (0)	97.50	89.17	5.00
	1 correct (1)	97.50	89.17	10.00
	2/3 correct (2)	97.50	89.17	23.33
	4/5 correct (3)	97.50	89.17	61.67
Language	Sentence repetition "cat"(1)	100	100	79.17
	Sentence repetition "John" (1)	100	99.17	61.34
	Fluency (1)	100	90.83	56.67
Abstraction	Similar train/bike (1)	100	100	65.83
	Similar watch/ruler (1)	99.17	99.17	75.00

Note. MoCA = Montreal Cognitive Assessment.
[a]"Complete response" and "response correct" were not recorded since the immediate recall item is not scored.
[b]Indicates the proportion of participants with a "complete response" for all of the delayed recall items.

clock rather than the cube into the CCFM illustrates this process of item selection. Time limitations necessitated that only one of the visuo-construction tasks could be included. Although the clock task had a greater median time required in the pre-test (1.22 min vs 0.97 min), we chose to keep it rather than the cube, because it offered an assessment of multiple components within the visuo-construction domain (contours, numbers, and hands) that are assessed as independent measures. The clock can also be used as a determinant of executive function and has a wealth of support in the literature on measure performance and other health outcomes. Pre-test data also indicated a lower completion response rate for the cube compared to the clock, 89.17% versus 92.50%, respectively. Moreover, field interviewer notes indicated respondent's dissatisfaction with drawing the cube. As a result, we decided that including the clock would offer greater participant acceptance and variation in response when evaluating cognition in the Wave 2 sample.

Table 5 compares the cognitive domains and individual items that are part of the SPMSQ, the MoCA, and CCFM. The CCFM items by order of presentation and ascribed points for a correct response are as follows (Table 5): orientation—"day" (1 point) and "month" (1 point); naming—"rhinoceros" (1 point); visuo-construction skills—"clock" (3 points); executive function—"trails b" (1 point); memory—"immediate five word recall" (0 points); attention—"serial 7s" (3 points); "forward digit span" (1 point); and "backward digit span" (1 point); language—"sentence repetition-cat" (1 point); verbal fluency—"F's" (1 point); abstraction—"similarity watch/ruler" (1 point); and memory—"delayed five word recall" (5 points). Incorrect responses to individual items indicate difficulty within a particular cognitive domain. Total CCFM scores range from 0 to 20 where higher scores indicate overall better cognitive function. Total scores can be used as a continuous outcome or to adjust for cognition in a multivariable analysis as a predictor.

Table 4. Distribution of Pre-Test MoCA Scores, *N* = 120

MoCA score	Frequency (percent)
30	4 (3.3)
29	5 (4.2)
28	9 (7.5)
27	12 (10.0)
26	16 (13.3)
25	12 (10.0)
24	10 (8.3)
23	13 (10.8)
22	2 (1.7)
21	4 (3.3)
20	5 (4.2)
19	9 (7.5)
18	6 (5.0)
17	8 (6.7)
16	1 (0.8)
14	2 (1.7)
12	1 (0.8)
7	1 (0.83)

Note. MoCA = Montreal Cognitive Assessment.

Table 6 displays the performance on the CCFM for NSHAP Wave 2 participants reported as percent correct separated by domain and item. Participants scored highest on the orientation domain with 91.2% correctly reporting date and month. Animal naming (Rhinoceros) was correctly completed by 83.6% of respondents. Executive function (trails b) and abstraction (word similarity watch/ruler) were correctly completed by about 58.8% and 59.5% of participants, respectively. Visuo-construction (clock draw) was correctly completed by 43.5% of participants with 1.8 % not completing any portion correct. On memory (delayed recall), 15.4% of participants could not recall any of the five words, while 18.9% were able to recall all five words. Within the attention domain (digits forward and backwards and serial 7s), 45.9% of participants correctly completed the domain with the highest performance on the digit forward item (88.4% correct) and lowest performance on serial 7s item (55.5% completely correct). The language domain (sentence repetition—cat and verbal fluency—F's) was correctly completed by 33.3% of participants, with the sentence repetition and verbal fluency item correctly completed by 61.6% and 47.9%, respectively.

The weighted mean CCFM score overall was 13.9 (*SE* 0.13). Men and women had mean CCFM score of 13.6 (*SE* 0.15) and 14.2 (*SE* 0.17), respectively. Persons age less than 65, 65–75, and 76 and older had CCFM scores of 15.4 (*SE* 0.17), 14.4 (*SE* 0.14), and 12.4 (*SE* 0.16), respectively. Based upon the pre-test data, CCFM scores (range 0–20 points) were highly correlated with MoCA scores (range 0–30 points), Pearson's *r* = .973. Linear regression modeling projects MoCA scores from CCFM scores using the equation MoCA = (1.14 × CCFM) + 6.83, which can be utilized to generate MoCA scores within NSHAP Wave 2.

DISCUSSION

We were able to successfully integrate the MoCA into the pre-test wave of NSHAP based on adaptations from the pilot testing for use by non-medical field interviewers using a CAPI administration. While the time to complete individual items varied substantially, each was almost always attempted and completed by participants. At the same time, there was substantial variability in the likelihood of a correct response to individual items. Using this information and applying a priori selection principles for items, we were able to develop the CCFM for NSHAP. This included representation from all eight cognitive domains of interest, variability in the likelihood of a correct response, integration into CAPI technology, administration by non-medically trained field interviewers, and average administration time of 12 min. We were also able to develop a robust estimate of projected MoCA scores based upon CCFM performance.

While neuropsychological testing remains the gold standard for assessing "pathological" cognitive changes in the older adult population, the interest and necessity to measure cognition as part of population-based studies is of increasing interest (Nathan, Wilkinson, Stammers, & Low, 2001). This is particularly relevant as cognitive assessments at the population-level remain biased, often being geographically isolated and heavily skewed by social and socio-economic factors. Several barriers to the integration of cognitive measures as part of a population-based survey have been highlighted by Herzog and Wallace in their capacity as Assets and Health Dynamics Among the Oldest Old (AHEAD) study investigators (Herzog & Wallace, 1997). First, a cognitive measure usually developed for the clinical setting would need to be adapted successfully for administration in the home setting. At the same time, the length and content of the measure should not overwhelm the participant, nor impact their willingness to complete other components of the survey. Finally, participants need to be willing to attempt and complete the cognitive measure, so that the response rate is high and truly representative of the population. Our pilot and pre-test data presented here addressed all of these concerns, which was further substantiated by NSHAP Wave 2 results.

U.S. longitudinal studies of aging have begun to integrate some cognitive measures as part of the interview. For example, The Health and Retirement Study (HRS), has contributed longitudinal and cross-sectional information on cognition since 1992, and initially assessed four aspects of cognitive—a memory test, an abstract reasoning test, self-rated cognitive functioning, and a "functioning in cognitively demanding activities of daily living" (Wallace & Herzog, 1995). More recently, HRS has expanded cognitive testing to include additional domains such as attention and verbal fluency as part of the Telephone Battery for Cognitive Status (TICS); however, visuo-construction and executive function have not been added. A sub-study

Table 5. Differences in Cognitive Domains Assessed Among the Short Portable Mental Status Questionnaire (SPMSQ), Montreal Cognitive Assessment (MoCA), and Chicago-Cognitive Function Measure (CCFM)

	Short Portable Mental Status Questionnaire (SPMSQ)	Montreal Cognitive Assessment (MoCA)	Chicago-Cognitive Function Measure (CCFM)
Domain			
Orientation	Day of week	Day of week	Day of week
	NA	Month	Month
	Date	Date	NA
	Place	Place	NA
	NA	Year	NA
	NA	City	NA
Naming	NA	Rhinoceros[a]	Rhinoceros[a]
	NA	Lion[a]	NA
	NA	Camel[a]	NA
Visuo-construction	NA	Clock draw[a]	Clock draw[a]
	NA	Copy cube[a]	NA
Executive Function	NA	Trails b[a]	Trails b[a]
Attention	Serial 3s	Serial 7s	Serial 7s
	NA	Forward digit span	Forward digit span
	NA	Backward digit span	Backward digit span
	NA	Vigilance "A" [a]	NA
Language	NA	Sentence repetition "cat"	Sentence repetition "cat"
	NA	Sentence repetition "John"	NA
	NA	Fluency[a]	Fluency[a]
Abstraction		Similarity watch/ruler	Similarity watch/ruler
		Similarity train/bike	NA
Memory	*Autobiographical*	*Delayed recall*	*Delayed recall*
	Age	Face	Face
	Birthplace	Velvet	Velvet
	Mother's Maiden Name	Church	Church
	Historical	Daisy	Daisy
	U.S. President	Red	Red
	Former U.S. President		

Note. NA = not assessed.
[a]Paper and pencil items not included in CAPI.

Table 6. NSHAP Wave II Items Correct, Weighted Distributions by Chicago Cognitive Function Measure, *n* = 3,377

Domain (score range)													
		Score correct (%)											
(Total range: 0–20)	Item	0	95% CI	1	95% CI	2	95% CI	3	95% CI	4	95% CI	5	95% CI
Orientation (0–2)	Month	2.2	1.5–2.9	97.8	97.1–98.5	—		—		—		—	
	Date	8.3	7.0–9.6	91.7	90.4–93	—		—		—		—	
	Orientation total (2-items)	1.7	1.1–2.3	7.1	6.0–8.2	91.2	89.9–92.5	—		—		—	
Naming (0–1)	Animal	16.4	14.3–18.5	83.6	81.5–85.7	—		—		—		—	
Executive function (0–1)	Trails	41.2	38.2–44.3	58.8	55.7–61.8	—		—		—		—	
Visuoconstruction (0–3)	Clock	1.8	1.2–2.4	17.2	15.3–19.1	37.5	35.1–39.8	43.5	41.0–46.0	—		—	
Memory (0–5)	Delayed Recall	15.4	13.6–17.2	8.1	6.8–9.3	16.4	14.3–18.6	19.7	17.9–21.5	21.5	19.8–23.2	18.9	16.8–21.0
Attention (0–5)	Forward digits	6	9.8–13.5	88.4	86.5–90.2	—		—		—		—	
	Backward digits	20.5	18.2–22.8	79.5	77.2–81.8	—		—		—		—	
	Subtractions	12.5	10.5–14.4	10.0	8.8–11.2	22.0	20.3–23.6	55.5	53.0–58.1	—		—	
	Attention total (3-items)	2.4	1.6–3.2	5.1	3.8–6.4	9.6	8.3–11.0	13.2	11.8–14.7	23.6	21.8–25.5	45.9	43.3–48.6
Language (0–2)	Sentence	38.4	34.1–42.6	61.6	57.4–65.9	—		—		—		—	
	Fluency	52.1	49.4–54.8	47.9	45.2–50.6	—		—		—		—	
	Language total (2-items)	23.8	20.7–26.8	42.9	40.3–45.5	33.3	30.3–36.4	—		—		—	
Abstraction (0–1)	Similarity	40.5	37.7–43.3	59.5	56.7–62.3	—		—		—		—	

Note. CI = confidence interval.

of HRS, the Aging, Demographics, and Memory Study (ADAMS), incorporated a comprehensive neuropsychological assessment of a selected group to better differentiate and understand normal cognition from cognitive impairment no dementia and dementia (Langa et al., 2005).

At the international level, studies of aging have also begun to incorporate cognitive measures. For instance, the English Longitudinal Study of Health and Aging (ELSA), a large study of community-dwelling individuals in the United Kingdom, assessed time orientation, immediate and delayed verbal recall, prospective memory, verbal fluency, numerical ability, cognitive speed, and attention (Llewellyn, Lang, Langa, & Huppert, 2008). The Canadian Study of Health and Aging (CSHA) incorporated the modified mini-mental status exam which is similar to the Folstein Mini-Mental Status Exam, but adds date and place of birth, animal naming, similarities, and a second delayed recall (Teng & Chui, 1987). The CSHA was designed to better understand pathologic changes in cognition (e.g., the development of dementia), rather than non-pathologic changes that concomitantly occur with aging (McDowell, Hill, & Lindsay, 2001). Importantly, the U.S. and international studies continue to overlook important contributors to cognitive functioning, namely visuo-construction skills and executive function

The development and incorporation of the CCFM into NSHAP Wave 2 allowed us to gather robust cognitive information from a nationally representative sample. When considering the use of the CCFM, we do not recommend dichotomizing the cognitive measure at a specific "threshold" or "cut-off" for "normal functioning" as this has not been definitively established in the literature. Also, we do not offer specific recommendations regarding the adjustment of CCFM scores based upon education level, given differences in published reports (e.g., variance in number of points to add for lower educational attainment) obtained from largely clinical samples, and because ours is a probability-based, non-clinical research sample.

Also, the CCFM builds upon the SPMSQ, the cognitive measure incorporated as part of NSHAP Wave 1. Researchers can leverage the overlap among domains between NSHAP waves and measures to examine cognition changes over time or as an approach to control for cognition between the two waves. The CCFM inclusion of frequently overlooked cognitive domains, such as visuo-construction, presents opportunities for investigators to better understand its relationship with health.

Investigators interested in the relationship of individual cognitive domains (e.g., executive function, memory, or attention) as an outcome or predictor of outcomes should give the following careful consideration. When multiple items are available within an individual domain, we recommend summing of all the items and not the use of an individual one, as this was how the measure was initially developed and validated. A paucity of research

with the CCFM (and MoCA) exists at the domain and the individual item level, compared with the gold standard of neuropsychological evaluation. Moreover, individual items that have been extensively studied, such as the clock draw, lack validity data using the scoring method for the CCFM (and MoCA). Lastly, the cognitive literature continues to evolve and many items developed to assess one domain may also provide insights into other domains. For example, the clock draw assesses visuo-construction, but it also provides important insights into executive functioning.

While our study has many strengths—including an extensive pilot phase and the use of a reliable and valid cognitive measure—several limitations should be considered. The NSHAP pre-test collected information from a small sample size in order to thoroughly evaluate the survey process itself, with the primary goal to establish the feasibility of item administration in the field by non-medically trained field interviewers. While the cohort is demographically similar to those included in NSHAP Wave 2, the pre-test is not comprised of a nationally representative sample, limiting its ability to be generalized. Moreover, participants recruited to participate in the pre-test were financially compensated and had previously successfully completed NSHAP Wave 1 and/or were motivated to participate in survey research. These sampling characteristics may contribute to the high rates of completion of the cognitive measures. The reliability and validity of the newly derived measure continues to be established. For example, the content validity of the CCFM appears similar to the MoCA and measures one underlying construct of general cognitive function, with evidence of a bi-factor model structure (paper under review). Also, the relationship between individual items and neuropsychological assessment for individual domains has not been definitively established. The CCFM has fewer items than the MoCA, so caveats to scoring should be considered. Finally, the pre-test included a wide age-range and cohort effects may contribute to differences in the time needed to complete individual cognitive items and their responses.

In conclusion, we were able to demonstrate the feasibility of integrating a more robust cognitive measure into NSHAP Wave 2. Using our guiding principles, we were able to use the information generated from both the pilot and pre-test samples to create a survey-adapted version of the MoCA, which we refer to as the CCFM. Benefits of the CCFM include an assessment of eight cognitive domains that can be administered by non-medically trained personnel in the field using CAPI technology in 12 min or less. The measures exhibit variability in response in a non-dementia population within each domain, so that meaningful comparisons between cognition, health, and social factors are possible.

REFERENCES

Alwin, D. F., & Hofer, S. M. (2011). Health and cognition in aging research. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *66*, i9–i16. http://dx.doi.org/10.1093/geronb/gbr051

Ashford, J. W. (2008). Screening for memory disorders, dementia and Alzheimers disease. *Journal of Aging & Health*, *4*, 399–432. http://dx.doi.org/10.2217/1745509X.4.4.399

Bernstein, I. H., Lacritz, L., Barlow, C. E., Weiner, M. F., & DeFina, L. F. (2011). Psychometric evaluation of the Montreal Cognitive Assessment (MoCA) in three diverse samples. *The Clinical Neuropsychologist*, *25*, 119–126. http://dx.doi.org/10.1080/13854046.2010.533196

Blackwell, T., Yaffe, K., Ancoli-Israel, S., Schneider, J. L., Cauley, J. A., Hillier, T. A., … Stone, K. L. (2006). Poor sleep is associated with impaired cognitive function in older women: The study of osteoporotic fractures. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*, *61*, 405–410. http://dx.doi.org/10.1093/gerona/61.4.405

Herzog, A. R., & Wallace, R. B. (1997). Measures of cognitive functioning in the AHEAD study. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *52*, 37–48. http://dx.doi.org/10.1093/geronb/52B.Special_Issue.37

Ishizaki, T., Yoshida, H., Suzuki, T., Watanabe, S., Niino, N., Ihara, K., …Imanaka, K. (2006). Effects of cognitive function on functional decline among community-dwelling non-disabled older Japanese. *Archives of Gerontology and Geriatrics*, *42*, 47–58. http://dx.doi.org/10.1016/j.archger.2005.06.001

Langa, K. M., Plassman, B. L., Wallace, R. B., Herzog, A. R., Heeringa, S. G., Ofstedal, M. B., … Willis, R. J. (2005). The aging, demographics, and memory study: Study design and methods. *Neuroepidemiology*, *25*, 181–191. http://dx.doi.org/10.1159/000087448

Llewellyn, D. J., Lang, I. A., Langa, K. M., & Huppert, F. A. (2008). Cognitive function and psychological well-being: Findings from a population-based cohort. *Age and Ageing*, *37*, 685–689. http://dx.doi.org/10.1093/ageing/afn194

Luis, C. A., Keegan, A. P., & Mullan, M. (2009). Cross validation of the Montreal Cognitive Assessment in community dwelling older adults residing in the Southeastern US. *International Journal of Geriatric Psychiatry*, *24*, 197–201. http://dx.doi.org/10.1002/gps.2101

McDowell, I., Hill, G., & Lindsay, J. (2001). An overview of the Canadian Study of Health and Aging. *International Psychogeriatrics*, *13*, 7–18. http://dx.doi.org/10.1017/S1041610202007949

Nasreddine, Z. (2003). Montreal Cognitive Assessment (MoCA). Retrieved from http://www.mocatest.org

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*, 695–699. http://dx.doi.org/10.1111/j.1532-5415.2005.53221.x

Nathan, J., Wilkinson, D., Stammers, S., & Low, J. L. (2001). The role of tests of frontal executive function in the detection of mild dementia. *International Journal of Geriatric Psychiatry*, *16*, 18–26. http://dx.doi.org/10.1002/1099-1166(200101)16:1<18::AID-GPS265>3.0.CO;2-W

Nazem, S., Siderowf, A. D., Duda, J. E., Have, T. T., Colcher, A., Horn, S. S., … Weintraub, D. (2009). Montreal Cognitive Assessment performance in patients with Parkinson's disease with "normal" global cognition according to mini-mental state examination score. *Journal of the American Geriatrics Society*, *57*, 304–308. http://dx.doi.org/10.1111/j.1532-5415.2008.02096.x

Paulo A. C., Sampaio, A., Santos, N. C., Costa, P. C., Cunha, P., Zihl, J., … Sousa, N. (2011). Patterns of cognitive performance in healthy ageing in Northern Portugal: A cross-sectional analysis. *PLoS ONE*, *6*, e24553. http://dx.doi.org/10.1371/journal.pone.0024553

Petersen, R. C., Smith, G. E., Waring, S. C., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1999). Mild cognitive impairment: clinical characterization and outcome. *Archives of Neurology*, *56*, 303–308.

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*, 433–441.

Potvin, O., Lorrain, D., Forget, H., Dubé, M., Grenier, S., Préville, M., & Hudon, C. (2012). Sleep quality and 1-year incident cognitive impairment in community-dwelling older adults. *Sleep*, *35*, 491–499. doi:10.5665/sleep.1732

Price C. C., Cunningham, H., Coronado, N., Freedland, A., Cosentino, S., Penney, D. L., … Libon, D. J. (2011). Clock drawing in the Montreal Cognitive Assessment: Recommendations for dementia assessment. *Dementia and Geriatric Cognitive Disorders*, *31*, 179–187. http://dx.doi.org/10.1159/000324639

Rossetti, H. C., Lacritz, L. H., Cullum, C. M., & Weiner, M. F. (2011). Normative data for the Montreal Cognitive Assessment (MoCA) in a population-based sample. *Neurology*, *77*, 1272–1275. doi:10.1212/WNL.0b013e318230208a

Salthouse, T. (2012). Consequences of age-related cognitive declines. *Annual Review of Psychology*, 2012, *63*, 201–226. http://dx.doi.org/10.1146/annurev-psych-120710-100328

Santos N. L., Costa P. S., Cunha P., Cotter, J., Sampaio A., Zihl, J., …, Sousa N. (2013). Mood is a key determinant of cognitive performance in community-dwelling older adults: A cross-sectional analysis. Age, 35, 1983–1993. doi:10.1007/s11357-012-9482-y

Steinerman, J. R., Hall, C. B., Sliwinski, M. J., & Lipton, R. B. (2010). Modeling cognitive trajectories within longitudinal studies: A focus

on older adults. *Journal of the American Geriatrics Society*, *58*, S313–S318. http://dx.doi.org/10.1111/j.1532-5415.2010.02982.x

Teng, E. L., & Chui, H. C. (1987). The Modified Mini-Mental State (3MS) examination. *The Journal of Clinical Psychiatry*, *48*, 314–318.

Wallace, R. B., & Herzog, A. R. (1995). Special issue on the Health and Retirement Study: Data quality and early results—Overview of the health measures in the Health and Retirement Study. *The Journal of Human Resources*, *30*, S84–S107. http://dx.doi.org/10.2307/146279