Published in final edited form as:

Alzheimer Dis Assoc Disord. 2009 ; 23(4): 384–388. doi:10.1097/WAD.0b013e31819e0bfc.

A Battery of Tests for Assessing Cognitive Function in Older Latino Persons

Kristin R. Krueger, PhD, Robert S. Wilson, PhD, David A. Bennett, MD, and Neelum T. Aggarwal, MD

Rush Alzheimer's Disease Center (KRK, RSW, DAB, NTA), and Departments of Behavioral Sciences (KRK, RSW), and Neurological Sciences (RSW, DAB, NTA), Rush University Medical Center, Chicago, IL, USA.

Abstract

With the proportion of older Latinos in the United States rapidly growing, dementia is expected to be an increasing public health problem in this segment of the population. Yet relatively few cognitive test batteries have been developed for evaluating older Spanish speaking persons. We selected a battery of cognitive tests used in cognitive aging studies of English speakers, adapted it for Spanish speakers, and administered it to 66 older Latinos (mean age = 71.1, SD = 8.1). The results of a factor analysis supported grouping the tests into the same five functional domains identified for English speakers. Composite measures of performance in each domain were positively related to education and, with some exceptions, inversely related to age. The results suggest that this battery may be useful in epidemiologic research on cognition in older Latinos.

Keywords

Spanish language tests; cognitive domains; cognitive function; longitudinal studies

Introduction

Decline in cognitive function is among the most feared consequences of old age. Recent research has shown, however, that cognitive decline is not an inevitable consequence of growing old. Although many older people do experience substantial cognitive decline, others experience slight decline, remain stable, or improve [1,2]. Increasing our understanding of aging and its consequences is therefore of utmost importance, as persons older than 65 years of age are the fastest growing segment of the United States population. The increase in Hispanic elders is projected to be especially sharp [3], yet our knowledge about age-related cognitive decline in Spanish speakers and the risk factors associated with such decline are limited.

Progress in understanding the consequences of aging among Latinos has been limited by a number of challenges. First, the majority of aging research has been done with non-Hispanic whites who differ from Latinos in important ways. For example, Fitten and colleagues [4] found that compared to data from white populations, Spanish speaking Latinos presented with different proportions of Alzheimer's disease and Vascular dementia and had a higher incidence of depression. Secondly, there has been a relative lack of tools that have been adapted for assessing cognitive decline in Spanish speakers, as well as incomplete knowledge of their psychometric properties. Thirdly, much research on aging among Latinos has been cross-

sectional and as a result, many adaptations of cognitive measures have focused on obtaining equivalent levels of performance of English and Spanish versions [5–8]. One concern in this approach is the difficulty of establishing equivalency of cognitive tests in different languages. It is furthermore debatable whether it is possible to identify truly comparable groups of Spanish-speaking and English-speaking elders, because Hispanic elders have substantially different cultural and educational backgrounds than non-Hispanic elders living in the United States.

Conversely, a longitudinal design addresses these challenges and offers the most direct way to assess cognitive aging. In a longitudinal study of cognition in older people, the most pressing concern is not that the original and adapted versions of a test have the same difficulty level, but rather that they each measure the same underlying construct. Outcome measures in such studies must be able to accommodate a wide range of performance since individual differences at baseline become more pronounced over time, as performance improves in some and declines in others. To further accommodate the wide range of performances, composite cognitive measure [2] are formed that combine several tests of the same ability but differing difficulty levels. Thus, a cognitive battery intended to measure change in different cognitive abilities needs multiple measures of each ability. The success of such an adaptation of a cognitive test battery can be determined by evidence that the tests are measuring the same underlying abilities in each language group.

In the present study, we adapted a battery of tests used to assess age-related cognitive decline in English speakers for use with older Spanish speakers. The original battery was chosen because it has been validated in two different cohorts of older English speakers [2,9]. The goal of adapting these tests was to obtain words that are universally understood, which is especially important for memory tests, which rely on a general familiarity of the words and expressions to carry out the task of memory evaluation. However, we also considered the importance of accounting for language that is region specific. We then administered the adapted Spanish battery to a group of older Spanish speakers and examined the extent to which an empirically based grouping of the tests based on factor analysis conformed to domains identified in English speakers. We hypothesized that the 18 individual tests could be grouped into the five functional domains based on previous research with similar sets of tests administered to English speakers [2], and would have the expected associations with age and education.

Methods

Participants

Participants were 66 community volunteers who met the inclusion criteria, which required them to be at least 55 years old, self-identify as a Spanish speaker and of Hispanic/Latino heritage, and not have a clinical diagnosis of dementia. Following a presentation about the study at various senior housing facilities and community centers, potential volunteers were asked to complete an interest form, which included questions about demographic information and level of interest in participating in the study. Those who expressed interest in the study were later contacted by project staff that provided participation requirements and obtained informed consent. Participants donated their time and information to the study without remuneration. The study was approved by the Institutional Review Board of Rush University Medical Center.

Participants had a mean age of 71.1 years ($SD = 8.1$), mean education was 6.9 years ($SD = 3.6$), and the mean score on the Mini-Mental State Examination [10] was 24.4 ($SD = 4.3$); 83% were women; 83% reported Mexican heritage, 15% reported Puerto Rican heritage, and 2% ($N=1$) reported Guatemalan heritage.

Assessment of Cognitive Function

We adapted a battery of 18 tests that has been used in previous longitudinal studies of English speakers [2] into Spanish. We chose this battery because composite measures of five cognitive domains have been created from it and have been used to document age-related change in cognitive function [2], including nonlinear patterns [2], with rate of change in different cognitive domains differentially related to risk factors [11], to outcomes like MCI [12], and death [13]. The individual tests in the original battery were chosen and/or modified so that they would be appropriate for elders of diverse backgrounds, many of whom had sensory and motor impairments. Specific modifications included enlarging the visual stimuli and in two cases reducing the number of test items administered as described below.

Seven tests in the battery assessed episodic memory: East Boston Story [2,14], Story A from Logical Memory of the Wechsler Memory Scale-Revised [15], and Word List Memory, Recall, and Recognition [16]. Semantic memory was assessed with two measures: the 15-item Consortium to Establish a Registry for Alzheimer's Disease (CERAD) version [16] of the Boston Naming Test [17], and verbal fluency [16]. Three tests of working memory were administered: Digit Span Forward and Digit Span Backward from the Wechsler Memory Scale-Revised [15] and Digit Ordering [18]. Perceptual speed was assessed with four measures: the oral version of the Symbol Digit Modalities Test [19], Number Comparison [20], and a modified version of the Stroop Neuropsychological Screening Test [21,22]. Visuospatial ability was tested using a 15-item form of Judgment of Line Orientation [23], and a 17-item form of Standard Progressive Matrices [24]. The Mini-Mental State Examination [10] was also administered and used to describe participants, but not used in analyses.

Adaptation of Measures into Spanish

We had two major goals in adapting the battery from English into Spanish. First, because the battery has been used in longitudinal studies for several years, we wanted to adapt the tests to be linguistically and culturally appropriate, while maintaining them as close to the original English version as possible. Such an adaptation would facilitate cross-cultural comparisons of data. Second, we wanted to use words and expressions that are universally understood by Spanish speakers regardless of region of origin.

Various methods for achieving a neutral or universally understood Spanish language have been discussed previously [5,25]. In our study we accomplished this using a combination of the back translation method [26] and group consensus [25], also referred to as the committee translation method [5]. First, the cognitive test battery was translated into Spanish and back translated into English. The back translation version and original battery were compared for inconsistencies. Second, a group of bilingual professionals reviewed the tests for comprehensibility, flow of language, and appropriateness of vocabulary choice. The consultants in this group were familiar with neuropsychology and research on aging, either due to their employee status in our center or as graduate students in neuropsychology. Additionally, several regions in Latin America, were represented in our consultant group including Mexico, the Caribbean, Central America and South America, facilitating the goal of eliminating words that are not universally understood [25].

It was necessary, albeit infrequently, to make minor adjustments for tests that probe abilities other than memory. For example, in the case of the Complex Ideational test that measures comprehension of language, we provided two choices for the word 'rubber' as in "Will water go through a good pair of rubber boots? (¿Se metería agua a un buen par de botas de hule/goma?) based on recommendations from the panel of language experts. Additionally, the accepted answer list for the Boston Naming Test is based on a review of existing translations [27] and piloting of this test, in order to avoid giving less common, but correct responses credit.

Both goals were met satisfactorily in the laboratory and were field tested on six predominantly Mexican and Puerto Rican non-neurological patients at an ambulatory clinic before testing the battery on a larger group of community participants. Minor changes were made to the battery following pilot testing and those changes are discussed in the results. The majority (95%) of the participants were interviewed by a bi-lingual Mexican-American research assistant. The remaining 5% were administered by a bi-lingual post-doctoral fellow who learned Spanish as a second language and had years of experience conducting evaluations in Spanish.

Data Analysis

We performed a principal-components factor analysis with varimax rotation on the 18 tests in the cognitive battery to test the hypothesis that the factor loadings of this battery were consistent with the conceptually based groupings. In order to assess the agreement of the conceptualized grouping with the empirical grouping, Rand's statistic [28] was used. The permutation test was used to test the likelihood that the fit between the empirically based and hypothesized groups could have been obtained by chance. Linear regression models controlled for sex were used to examine the associations of age and education with the composite measures of each cognitive domain. Programming was done in SAS [29].

Results

The mean and standard deviation of each of the 18 cognitive function tests used in analyses are shown in Table 1. One test, Word List Recognition, had a negatively skewed distribution. The remaining 17 tests had approximately symmetric distributions.

We tested the hypothesis that the 18 individual tests could be grouped into the five functional domains shown in Table 1 based on previous research with similar sets of tests administered to English speakers [2]. To test this hypothesis, we developed an empirically based grouping by performing a factor analysis of the 18 tests with a varimax rotation. As shown on the right side of Table 1, five factors emerged from this analysis. We then grouped tests with loading of 0.5 or greater, to be consistent with previous research [2], on a common factor and used Rand's statistic [28] to assess the goodness of fit between the hypothesized and empirically obtained groupings. For this analysis, we rescaled Rand's statistic to range from -1 (complete disagreement) between the two groupings to +1 (complete agreement between the two groupings) to be comparable to Kendall's tau, another measure that calculates proportion of pairs concordant minus proportion discordant, as previously described [2]. Rand's statistic was 0.73 ($p < 0.001$), indicating a good fit between the factor analytic results and the hypothesized grouping.

Summary measures of each of the five domains were formed as in previous research [2] by converting raw scores on each component test to z scores, using the mean and standard deviation of all participants, and then averaging the z scores of the component tests in that domain. If more than half of the component tests were missing, the domain score was treated as missing. Table 2 lists the mean, standard deviation, and range for the five cognitive domain scores. None of the participants achieved the minimum or the maximum possible composite score on any of the five composite domains measures, suggesting that they can accommodate wide individual differences in ability. By way of comparison, minimum scores were obtained on 10 of the 18 individual tests; maximum scores were obtained on 5 of the 18 tests both of these on 3 of the 18 individual tests.

In a final set of linear regression analyses, we regressed each composite measure on age, sex, and education to test whether performance in each cognitive domain measure would be inversely related to age and positively related to education. As shown in Table 3, education was positively related to performance in all five domains. Age was inversely related to episodic

memory, semantic memory, and perceptual speed; however, the associations of age with working memory and visuospatial ability approached, but did not meet statistical significance.

Discussion

We adapted a battery of tests designed to measure change in cognitive function in English speakers for use with Spanish speakers. We used a consensus method of translation in order to create a Spanish version that could be used with Spanish speakers in the United States, regardless of country of origin. We then administered this battery to Spanish speaking community participants. In a factor analysis, five factors emerged which corresponded well with five hypothesized cognitive domains found in English speakers. As expected, performance in each cognitive domain was positively related to education and, with some exceptions, inversely related to age. Overall, the results suggest that the original battery and adaptation provide conceptually similar measures of cognitive functioning, thereby making the psychometric properties available for a battery that may be useful in assessing age-related change in cognitive function in older Spanish speakers.

In this study we tested the concordance between a hypothesized grouping of the individual tests and an empirically-based grouping using Rand's statistic. The obtained value, 0.73, suggested a moderately good fit. Very similar versions of this test battery have been administered to English speakers, and factor analytically-based test groupings have been compared to the same five groups hypothesized here, using Rand's statistic to quantify goodness-of-fit. In a cohort of older Catholic clergy members, Rand's statistic was 0.79 [2]. In the Rush Memory and Aging project, Rand's statistic was 0.57 when analyzed in the first 141 participants [9] and was 0.79 when data on more than 500 persons were analyzed [22]. That the fit between hypothesized and obtained test groupings in this group of Spanish speaking older persons is so similar to the fits obtained in previous studies of English speakers supports the idea that the test battery is assessing the same underlying cognitive abilities in English and Spanish speakers, consistent with a previous study in which the results of an exploratory factor analysis of an adapted test battery were judged to be consistent with previous research on it in English [30].

It is difficult to compare the level of cognitive performance in people of different cultural backgrounds because the influences of culture are imperfectly understood and measured. Although the association of culture with level of cognitive performance is difficult to isolate, it is probably true that cultural influences are relatively constant over time, at least for the brief temporal intervals covered in most longitudinal aging studies. As a result, change in cognitive function should be relatively free of these biases and so offers a powerful way to compare aging across ethnic subgroups. The present cross-sectional result suggests that this battery has two important features needed in cognitive test batteries to be used in comparing subgroups. The first feature is that the test battery appears to measure the same underlying constructs in English. Secondly, the constructs are measured with multiple tests, which are combined into composite measures. This method of measurement is able to reflect a much wider range of performance of the sort needed to capture change in older people of initially different ability levels. Longitudinal research will be needed to confirm the properties of this battery in a larger group of people.

This study has important limitations. The battery was tested on a small selected group of community residents, and therefore it is unlikely that the full range of cognitive activity has been represented. Due to the restricted range of education, we may have tested the lower ranges of performance, but most likely did not test the upper ranges. The majority of the participants in our study were Mexican women, and the results may not extend to other Spanish speaking populations due to cultural, linguistic and gender differences. Additionally, we did not assess

other important variables, such as bilingual status, level of acculturation, country where education was completed and number of years in the United States, which might have affected results. Further research on this tests battery with a larger, more diverse group of participants is needed to more securely establish its psychometric properties.

Acknowledgments

This project was supported by NIA center grant P30 AG10161. We are grateful to our participants who were recruited from several different communities in Chicago and from the following senior housing buildings and community organizations: Villa Guadalupe, South Chicago YMCA, West Town/Logan Square, Pilsen, Hispanic Housing Association, Centro Comunitario Juan Diego. We also thank Karen Graham, MA and Mary Futrell for recruitment efforts; Liping Gu, for statistical programming, and George Dombrowski, MS and Greg Klein for data management.

The research was supported by grant P30 AG10161 from the National Institute on Aging.

References

1. Christensen H, MacKinnon AJ, Korten AE, Jorm AF, et al. An analysis of diversity in the cognitive performance of elderly community dwellers: Individual differences in change scores as a function of age. *Psychol Aging* 1999;14:365–379. [PubMed: 10509693]
2. Wilson RS, Beckett LA, Barnes LL, et al. Individual differences in rates of change in cognitive abilities of older persons. *Psychol Aging* 2002;17:179–193. [PubMed: 12061405]
3. Hobbs, F. The UNITED STATES Census Bureau 2001. The elderly population. Online: www.census.gov/population/www/pop-profile/elderpop.html
4. Fitten LJ, Ortiz F, Ponton M: Frequency of Alzheimer's disease and other dementias in a community outreach sample of Hispanics. *J Am Geriatr Soc* 2001;1301–1308. [PubMed: 11890488]
5. Loewenstein DA, Arguelles T, Barker WW, Duara R. A comparative analysis of neuropsychological test performance of Spanish-speaking and English-speaking patients with Alzheimer's disease. *J Gerontol B Psychol Sci Soc Sci* 1993;48:P142–P149.
6. Taussig IM, Henderson VW, Mack W. Spanish translation and validation of a neuropsychological battery: Performance of Spanish-and English-speaking Alzheimer's disease patients and normal comparison subjects. *Clin Gerontol* 1992;2:95–107.
7. Mungas D, Reed BR, Marshall SC, Gonzalez HM. Development of psychometrically matched English and Spanish language neuropsychological test for older persons. *Neuropsychology* 2000;14:209–223. [PubMed: 10791861]
8. Stricks L, Pittman J, Jacobs DM, Sano M, Stern Y. Normative data for a brief neuropsychological battery administered to English-and Spanish-speaking community-dwelling elders. *JINS* 1998;4:311–318. [PubMed: 9656604]
9. Wilson RS, Barnes LL, Bennett DA. Assessment of lifetime participation in cognitively stimulating activities. *JCEN* 2003;25:632–642.
10. Folstein MF, Folstein SE, McHugh PR. Mini-Mental State: A practical method for grading the mental state of patients for the clinician. *J Psychiatri Res* 1975;12:189–198.
11. Wilson, Schneider, Barnes, et al. The apolipoprotein E E4 allele and decline in different cognitive systems during a 6-year period. *Arch Neurol* 2002;1154–1160. [PubMed: 12117364]
12. Bennett DA, Wilson RS, Schneider JA. natural history of mild cognitive impairment in older persons. *Neurology* 2002;59:198–205. [PubMed: 12136057]
13. Wilson RS, Beckett LA, Bienias JL, et al. Terminal decline in cognitive function. *Neurology* 2002;1154–1160. [PubMed: 11971080]
14. Albert M, Smith L, Scherr P, Taylor J, Evans D, Funkenstein H. Use of brief cognitive tests to identify individuals in the community with clinically diagnosed Alzheimer's disease. *Int J Neurosci* 1991;57:167–178. [PubMed: 1938160]
15. Wechsler, D. Wechsler Memory Scale-Revised Manual. San Antonio: Psychological Corp.; 1987.
16. Welsch KA, Butters N, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). V. A normative study of the neuropsychological battery. *Neurology* 1994;44:609–614. [PubMed: 8164812]

17. Kaplan, EF.; Goodglass, H.; Weintraub, S. The Boston Naming Test. Vol. ed2. Philadelphia: Lea & Febiger; 1983.
18. Cooper JA, Sager HJ, Jordan N, et al. Cognitive impairment in early, untreated Parkinson's disease and its relationship to motor disability. *Brain* 1991;114:2095–2122. [PubMed: 1933236]
19. Smith, A. Symbol Digit Modalities Test Manual, revised. Los Angeles: Western Psychological Services; 1982.
20. Ekstrom, RB.; French, JW.; Harman, HH., et al. Manual for factor-referenced cognitive tests. Princeton: Educational Testing Service; 1976.
21. Trenerry, MR.; Crosson, B.; DeBoe, J., et al. The Stroop neuropsychological screening test. Odessa, FL: Psychological Assessment Resources; 1989.
22. Wilson RS, Barnes LL, Krueger KR, Hoganson G, Bienias JL, Bennett DA. Early and late life cognitive activity and cognitive systems in old age. *JINS* 2005;11:400–407. [PubMed: 16209420]
23. Benton, AL.; Sivan, AB.; Hamsher, K., et al. Contributions to neuropsychological assessment ed2. New York: Oxford University Press; 1994.
24. Raven, JC.; Court, JH.; Raven, J. Manual for Raven's progressive matrices and vocabulary: Standard Progressive Matrices. Oxford: Oxford Psychologists Press; 1992.
25. Woodcock, RW.; Munoz-Sandoval, AF. The batería in neuropsychological assessment. In: Pontón, MO.; León-Carrión, J., editors. *Neuropsychology and the Hispanic Patient: A Clinical Handbook*. Mahwah, NJ: Lawrence Erlbaum Associates; 2001. p. 137-164.
26. Brislin RW. Back-translation for cross-cultural research. *J Cross Cult Psychol* 1970;1:185–216.
27. Pontón MS, Satz P, Herrera L, et al. Modified Spanish version of the Boston Naming Test. *Clin Neuropsychol* 1992;3:334.
28. Rand WM. Objective criteria for the evaluation of clustering methods. *J Am Stat Assoc* 1971;66:846–850.
29. SAS Institute Inc. SAS OnlineDoc® 9.1.3. Cary, NC: SAS Institute Inc; 2004.
30. Pontón MO, Gonzalez JJ, Hernandez I, Herrera L, Higareda I. Factor analysis of the Neuropsychological Screening Battery for Hispanics (NeSBHIS). *Appl Neuropsychol* 2000;7:32–39. [PubMed: 10800626]

Table 1

Psychometric Information on the 18 Cognitive Function Tests

<i>Cognitive Test</i>	<i>Mean</i>	<i>SD</i>	<i>Skew</i>	<i>Hypothesized Domain</i>	Rotated Factor Loading				
					1	2	3	4	5
Logical Memory Ia	8.2	4.4	0.4	Episodic memory	0.72	0.50	-0.04	0.06	0.07
Logical Memory Iia	6.6	4.4	1.0	Episodic memory	0.74	0.41	0.15	0.10	0.02
Immediate story recall	8.5	2.0	-0.2	Episodic memory	0.82	-0.08	-0.00	0.07	0.06
Delayed story recall	7.4	2.9	-1.1	Episodic memory	0.86	-0.09	0.17	0.12	-0.02
Word List Memory	13.1	4.5	0.0	Episodic memory	0.55	0.14	0.27	0.21	0.46
Word List Recall	4.0	2.4	-0.2	Episodic memory	0.66	0.24	0.20	-0.28	0.33
Word List Recognition	8.9	2.1	-2.6	Episodic memory	0.19	0.02	0.66	0.18	-0.21
Boston Naming Test	12.1	2.3	-0.7	Semantic memory	0.16	0.64	0.48	-0.01	0.18
Category Verbal Fluency	30.5	7.7	0.1	Semantic memory	0.28	0.34	0.16	0.09	0.58
Digit Span Forward	4.7	1.9	0.1	Working memory	-0.14	0.12	-0.12	0.74	0.28
Digit Span Backward	3.9	1.8	0.8	Working memory	0.24	0.31	0.34	0.58	-0.05
Digit Ordering	4.5	2.5	-0.3	Working memory	0.24	0.46	0.23	0.46	0.33
Symbol Digit Modalities Test	23.3	14.0	0.0	Perceptual speed	0.06	0.28	0.72	0.07	0.15
Number Comparison	13.3	7.7	0.5	Perceptual speed	0.03	0.18	0.80	-0.00	0.24
Stroop Color Naming	15.0	7.3	-0.3	Perceptual speed	0.16	-0.05	0.38	0.63	0.32
Stroop Word Reading	40.0	14.3	0.0	Perceptual speed	0.00	0.02	0.01	0.29	0.80
Judgment of Line Orientation	8.3	3.7	-0.5	Visuospatial ability	-0.01	0.78	0.20	0.17	0.10
Standard Progressive Matrices	8.7	2.5	-0.7	Visuospatial ability	0.22	0.63	0.06	0.59	0.02

Table 2

Psychometric Information on the Cognitive Domain Measures

Cognitive domain	<i>M</i>	<i>SD</i>	Range
Global cognition score	0.084	0.644	−1.449 to 1.410
Episodic memory	0.107	0.733	−1.648 to 1.896
Semantic memory	0.110	0.759	−1.777 to 1.363
Working memory	0.052	0.807	−1.543 to 2.208
Perceptual speed	0.115	0.747	−1.426 to 1.698
<i>Visuospatial ability</i>	0.026	0.923	−2.395 to 1.599

Note Each cognitive domain measure is the mean of the *z* scores on individual tests grouped in the domain.

Table 3

Relation of age and education to composite measures of function in different cognitive domains*

<i>Cognitive Domain</i>	<i>Model Term</i>	<i>Estimated Slope</i>	<i>(SE)</i>	<i>p</i>
Episodic memory	Age	−0.033	0.010	0.002
	Education	0.071	0.022	0.002
Semantic memory	Age	−0.030	0.012	0.013
	Education	0.078	0.025	0.002
Working memory	Age	−0.160	0.012	0.198
	Education	0.057	0.028	0.042
Perceptual speed	Age	−0.026	0.012	0.026
	Education	0.064	0.025	0.013
Visuospatial ability	Age	−0.022	0.292	0.128
	Education	0.080	0.031	0.011

*
From separate linear regression models adjusted for sex.