Published in final edited form as:

Appl Neuropsychol Adult. 2015 ; 22(6): 435–444. doi:10.1080/23279095.2014.978451.

Neuropsychological Language Tests in Dementia Diagnosis in English-Speaking Hispanic and Non-Hispanic White Outpatients

Philip Sayegh

Department of Psychiatry and Biobehavioral Sciences, University of California, Los Angeles, David Geffen School of Medicine, Los Angeles, California.

Abstract

Neuropsychological language tests have limitations (e.g., unrepresentative norms) when applied to “Hispanics” of which clinicians are likely aware that may lead to inaccurate dementia diagnoses. Therefore, it was hypothesized that language tests would be weaker diagnostic predictors in Hispanics versus “non-Hispanic Whites.” Participants included 436 English-speaking Hispanic and 436 non-Hispanic White (randomly selected from 10,937) outpatients classified as *dementia* or *normal cognition* at initial evaluation. When covarying for age, sex, education, and functional abilities, vegetable fluency significantly predicted diagnosis among non-Hispanic Whites, odds ratio = 0.80, 95% confidence interval [0.69, 0.94], $p < .01$. Animal fluency and an abbreviated (30-item) Boston Naming Test (BNT) comparably predicted diagnosis across groups. Results remained similar when covarying for primary language among Hispanics. Findings suggest that vegetable but not animal fluency was relatively unimportant in diagnosis for the English-speaking Hispanics in this study possibly because of cultural influences on the familiarity, salience, and relevance of this category’s items. Additionally, clinicians may have informally adjusted Hispanics’ 30-item BNT and animal fluency scores and discounted vegetable fluency to account for their limitations. Animal fluency and BNT may be preferable language tests when assessing dementia in English across groups, as they comparably predicted diagnosis in both groups.

Keywords

ethnicity; culture; Latino/a; Boston Naming Test; verbal fluency

The population of older Hispanics in the United States (US) with dementia is anticipated to grow considerably from less than 200,000 in 2000 to as much as 1.3 million by 2050 (Alzheimer’s Association, 2004). The label “Hispanic” was developed by the U.S. Census to classify individuals of Spanish-speaking Latin American descent, which embodies a total of 21 different nations (Suárez-Orozco & Pérez, 2002). Thus, Hispanics represent a very ethnically and racially diverse group. However, in spite of this within-group diversity, Valle and Lee (2002) predicted that elderly Hispanic Americans will face a disproportionate risk of dementia during approximately the upcoming 50 years for various reasons including

genetic and socioeconomic factors, highlighting the urgency of better comprehending ethnic-group differences in the assessment and diagnosis of dementia.

Although neuropsychological language tests are often included in dementia assessment, they consist of numerous limitations when applied in English to English-speaking Hispanics, especially those for whom English may not be their native or primary language. For example, the norms used for most language tests were developed using monolingual English-speaking Caucasian samples, calling the validity of their findings among many English-speaking Hispanic Americans into question (Portocarrero, Burright, & Donovick, 2007). Furthermore, potentially important factors, such as number of years living in the US, primary language, years and quality of education, and levels of acculturation, English-language proficiency, and literacy are often not considered when developing norms (Manly & Espino, 2004). English-language literacy in particular has been shown to have a strong effect on neuropsychological test performance (conducted in English) in Hispanic elders and is a more sensitive predictor of baseline performance beyond other key demographic and experiential factors (Manly, Byrd, Touradji, Sanchez, & Stern, 2004). Therefore, language tests may have reduced validity such that English-speaking Hispanics with normal cognitive function may be more likely to be inappropriately diagnosed with dementia compared to non-Hispanic Whites (e.g., Le Carret et al., 2003).

Indeed, several studies have reported that bilingual Spanish-English speakers perform significantly worse on tasks of semantic fluency than monolingual English speakers when assessed in English (e.g., Gollan, Montoya, & Werner, 2002; Rosselli et al., 2000) for several possible reasons. First, cross-language interference can contribute to the misdiagnosis of dementia. Hernandez and Kohnert (1999) suggested that older adults may be especially susceptible to cross-language interference, thereby inflating their chances of misdiagnosis. Second, Burke, MacKay, Worthley, and Wade (1991) suggested that tip-of-the-tongue states (i.e., word-retrieval failures in which speakers sense imminent recall and are capable of reporting features about the target word) may hinder performance on semantic fluency tasks in bilinguals. Notably, such tip-of-the-tongue states have been found to be more common in English-dominant bilinguals (i.e., not just Spanish-dominant bilinguals) compared to English-speaking monolinguals (Gollan & Silverberg, 2001). Finally, Gollan et al (2002) noted that verbal fluency performance among bilinguals (compared to monolinguals) may be reduced as they report being unfamiliar with very low-frequency target words. Such factors may ultimately result in reduced accuracy of dementia assessment and diagnosis among bilingual Hispanics.

Regarding confrontation naming tests (e.g., the Boston Naming Test [BNT]), studies have found that bilinguals named pictures more slowly than monolinguals when asked to provide responses in either their dominant or nondominant languages (e.g., Ivanova & Costa, 2008; Roberts, Garcia, Desrochers, & Hernandez, 2002). Therefore, bilinguals may be at a disadvantage on timed tasks of confrontation naming compared to monolingual English speakers. Consistent with this notion, prior research has shown that Spanish-English bilingual adults scored well below monolingual English-speaking adults on BNT when assessed in English (e.g., Kohnert, Hernandez, & Bates, 1998; Roberts et al., 2002), thereby limiting the validity and utility of this test in this group (Strauss, Sherman, & Spreen, 2006).

Clinicians may be aware of such limitations of language tests when applied in English to English-speaking Hispanics. Thus, they may interpret scores too leniently or over-adjust based on linguistic, cultural, or other demographic variables (Marquez de la Plata et al., 2009). Bortnik et al. (2013) also reported that the majority (58.6%) of clinical neuropsychologists surveyed in their study made adjustments for primary language abilities (i.e., English as a second language) when reporting BNT scores. Similarly, 46.9% made such adjustments based on patient ethnicity.

In a study involving a similar sample of English-speaking Hispanic and non-Hispanic White outpatients, Sayegh and Knight (2013a) found that overall neuropsychological test performance predicted a dementia diagnosis or a normal cognitive function classification to similar degrees across ethnic groups when using either combined-ethnic-group norms or ethnic-group-specific norms. They posited that clinicians may have weighed certain tests such as language tests differentially across groups to informally account for their reduced validity, though they did not formally test this supposition. A richer understanding of how clinicians may weigh language tests in particular to varying degrees across ethnic groups could shed light on possible systematic differences in dementia assessment and diagnosis.

The key objective of this study was to determine whether neuropsychological language tests differentially predicted diagnosis across English-speaking Hispanics and non-Hispanic Whites. It was hypothesized that language tests (i.e., of verbal fluency and confrontation naming) would be significantly stronger predictors of a classification of normal cognitive function or a dementia diagnosis in English-speaking non-Hispanic Whites than in their Hispanic counterparts.

Methods

Participants and Procedure

This research was compliant with institutional research standards for human research. The study population was composed of outpatients (and their informants) who provided informed consent to participate in the longitudinal National Alzheimer's Coordinating Centers Alzheimer's Disease Research Center study (Uniform Data Set) at 32 centers across the US. Data for the Hispanic participants in the current study were derived from the available 436 English-speaking Hispanic outpatients in the Uniform Data Set who met inclusion and exclusion criteria for this study (see below). For the non-Hispanic Whites, 436 English-speaking participants were randomly selected out of the available 10,937 non-Hispanic White outpatients in the Uniform Data Set using the *Select Cases* function in SPSS 17.0 to both decrease the likelihood of significant findings as a result of high statistical power and augment comparability across ethnicities. Inclusion criteria included being diagnosed with dementia or classified as having normal cognitive function at initial study evaluation. Exclusion criteria included: diagnoses of either Parkinson's disease or stroke without dementia, mild cognitive impairment (due to the small sample of patients with this diagnosis), and cognitive impairment without mild cognitive impairment or dementia; limited English-language proficiency (based on the fact that patients completed the Uniform Data Set using the English rather than the available Spanish module as deemed appropriate

by study personnel and based on participants' self-report); and no reliable informant available to provide information on patients' symptoms at initial evaluation.

The complete Uniform Data Set neuropsychological assessment battery, conducted in English, lasted approximately 1 hour. Individuals who administered the cognitive battery differed across sites and may have included physicians, neuropsychologists, neuropsychology technicians, research coordinators, and nurses. However, all sites followed guidelines for standardization of test administration and scoring as outlined in a detailed manual of administration and scoring instructions and other key study protocol procedures (Weintraub et al., 2009). Informants answered several questions about patients' functional abilities.

Measures

Demographics and covariates—Demographic variables contained in the Uniform Data Set include ethnicity, age, sex, and years of obtained education. These variables (excluding ethnicity) were included as covariates in the analyses, as they may affect both an individual's neuropsychological test performance and chances of cognitive decline. Ethnicity (Hispanic or non-Hispanic White) was used to classify the ethnic groups and stratify analyses. The regression analysis for the English-speaking Hispanics was also run with the inclusion of primary language (English or another language) as a covariate, and results were compared to the first model without this variable. As there was no formal measure of primary language available as part of this dataset, the classification of English versus Spanish as patients' primary language was based on the fact that patients completed the Uniform Data Set using the English rather than the available Spanish module as deemed appropriate by study personnel and based on participants' self-report. Primary language was not included in the non-Hispanic White model, as nearly all (98.62%; $n = 430$) reported that English was their primary language.

In addition, the Functional Assessment Questionnaire (FAQ; Pfeffer, Kurosaki, Harrah, Chance, & Filos, 1982) was included as a covariate in the analyses. The FAQ is a widely used scale that quantifies patients' ability to carry out instrumental activities of daily living (e.g., shopping and managing finances) to aid clinicians in the assessment and diagnosis of dementia. Clinicians use information provided directly from informants to complete this questionnaire. The FAQ measures a total of 10 instrumental activities of daily living and uses a rating scale ranging from *normal* to *dependent*. Overall scores for the FAQ span from 0 to 30, with higher scores indicating increased difficulty or needing support with instrumental activities of daily living over the prior 4 weeks. The FAQ was first validated on a sample of 195 older adults between the ages 61 to 91 years residing in a stable retirement community. Sayegh and Knight (2013b) recently confirmed the ethnic-group configural (i.e., number [1] of factors) and factorial (i.e., pattern of factor loadings) invariance of this scale using a similar sample of English-speaking Hispanic and non-Hispanic White outpatients. The Cronbach's alpha values for the FAQ in this sample were .97 for Hispanics and .96 for non-Hispanic Whites.

The FAQ was included as a covariate in all analyses, as functional abilities tap one of the two key criteria of a dementia diagnosis (i.e., declines in cognition and functioning;

American Psychiatric Association, 2000). As such, covarying this important diagnostic variable allows for a clearer picture of the role of neuropsychological language tests in the diagnostic process above and beyond this variable across ethnic groups.

Neuropsychological language tests—Three language tests were included in the Uniform Data Set that assessed two aspects of language. First, semantic (i.e., category) fluency was assessed by two tests that required patients to orally generate as many words within the categories of animals and vegetables as they could, with 60 seconds allotted for each test. Second, the 30 odd-numbered items of BNT (Kaplan, Goodglass, & Weintraub, 1983) were used as a measure of confrontation naming. Only the 30 odd-numbered BNT items were included in this larger comprehensive study's battery in the interest of brevity. Nonetheless, in a sample of individuals with probable Alzheimer's disease, other forms of dementia, and normal controls, this 30-item version has been shown to: (1) have good internal-consistency reliability, (2) significantly and highly correlate with the full version of BNT, and (3) allow for easy and relatively accurate extrapolation to a total BNT score given the lack of significant mean differences between versions across the three diagnostic groups (Williams, Mack, & Henderson, 1989). This test requires participants to name line drawings of objects within 20 seconds. Total scores were calculated as the sum of the number of correct spontaneous responses and the number correct following semantic (but not phonemic) cues. The Uniform Data Set also contained other neuropsychological tests that, as previously described, were examined in a prior study by Sayegh and Knight (2013a) in combination (i.e., using standardized overall sum of test scores) in terms of their role in diagnosis across ethnicities. These tests were: Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975), Logical Memory (Story A) subtest of the Wechsler Memory Scale—Revised (Wechsler, 1987a), Digit Span Forward and Backward and Digit Symbol Coding subtests of the Wechsler Adult Intelligence Scale—Revised (Wechsler, 1987b), and Trail Making Test Parts A and B (Reitan, 1958).

Clinical diagnosis of cognitive function—Clinical diagnoses were assigned, either individually by medical physicians (e.g., neurologists and psychiatrists) or through consensus with additional clinicians (e.g., neuropsychologists and nurses), using all Uniform Data Set information accessible for each patient. Clinicians provided either *yes* or *no* responses to questions on whether the patient: (1) had normal cognitive function (i.e., no dementia, mild cognitive impairment, or other neurological condition causing cognitive impairment); (2) had dementia in agreement with standard criteria for Alzheimer's disease (i.e., the National Institute of Neurological and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association [i.e., NINCDS/ADRDA] Alzheimer's criteria; McKhann et al., 1984); (3) had vascular dementia (i.e., based on the National Institute of Neurological Disorders and Stroke and Association Internationale pour la Recherche et l'Enseignement en Neurosciences [i.e., NINDS/AIREN] vascular dementia criteria; Román et al., 1993); or (4) demonstrated adequate evidence of dementias other than Alzheimer's or vascular dementia.

Statistical Method

To test the key hypotheses, multivariate binary logistic regression analyses were run stratified by ethnic group, with age, sex, education level, primary language (for the English-speaking Hispanics only to assess for any potential difference in the results among this ethnic group), and total FAQ scores included as covariates and the three language tests as the key predictor variables. As nearly all non-Hispanic Whites in this study reported that English was their primary language, this variable was not included as a covariate for this group and, thus, analyses were stratified by ethnic group. All continuous predictor variables were centered on their sample means to increase the interpretability to their intercepts and assist with avoiding multicollinearity. Multicollinearity was assessed by examining tolerance and variance inflation factor values, with tolerance values of $<.40$ and variance inflation factor values of >2.5 suggestive of multicollinearity (Allison, 2012). Odds ratios (*ORs*) for English-speaking Hispanics and non-Hispanic Whites were obtained for each of the three language tests in terms of their associations with classification or diagnosis, which was coded as a binary variable (0 = normal cognitive function; 1 = dementia).

As the Uniform Data Set is composed of data collected at multiple sites nationwide, variations across sites in terms of assessment and diagnosis procedures as well as other factors, such as patient characteristics and the distribution of diagnoses, were likely present. Therefore, logistic regression analyses were conducted using the Generalized Estimating Equations marginal models method with the GENMOD procedure in SAS 9.2, which employs a robust covariance matrix to account for correlated measurements within sites (Agresti, 2007). This procedure takes into consideration the possible confounding role of site and allows for the interpretation of parameter estimates to be independent of the respective site. In other words, it is suitable for the entire population of sites and in fact averages the effects of the independent variables across sites. To test whether there were significant differences in the *ORs* for each independent variable across ethnicities, the statistical significance of the *ORs* across groups was assessed and tests for interaction by ethnicity were conducted (Altman & Bland, 2003).

Results

Descriptive statistics for the demographic and key variables, separated by both ethnicity and classification or diagnosis (combined normal cognitive function and dementia, normal cognitive function only, and dementia only), are provided in Table 1.

Of the 436 English-speaking Hispanic outpatients in this sample, 232 (53.21%) were Mexican, Chicano, or Mexican American, 81 (18.58%) were Puerto Rican, 26 (5.96%) were South American, 19 (4.36%) were Central American, 19 (4.36%) were Cuban, and 8 (1.83%) were Dominican, with 30 (6.88%) were classified as *other* and 21 (4.82%) as *unknown*. Additionally, 297 (68.12%) indicated at the time of evaluation that English was their primary language. A total of 237 (54.36%) were assigned a diagnosis of dementia at their initial evaluations and the remainder classified as normal cognitive function. Of the 436 non-Hispanic White outpatients, almost all (98.62%; $n = 430$) reported English as their primary language. The percentage (53.90%; $n = 235$) of non-Hispanic Whites who were assigned a dementia diagnosis rather than classified as normal cognitive function was

similar to that of the Hispanics. Among all patients who were assigned a dementia diagnosis, the most common contributing factor to dementia was probable or possible Alzheimer's disease ($n = 384$, 81.36%), followed by frontotemporal dementia ($n = 40$, 8.47%) and dementia with Lewy bodies ($n = 33$, 6.99%). The frequencies of these contributing factors to dementia did not differ significantly across ethnic groups except for frontotemporal dementia, $\chi^2(1, n = 472) = 7.14, p = .01$, with a greater presence in non-Hispanic Whites ($n = 28$) compared to Hispanics ($n = 12$).

Table 1 also reports whether there were any statistically significant differences among the demographic and key variables. Among English-speaking Hispanics, patients classified as having normal cognitive function versus dementia were significantly more likely to have reported that English was their primary language. There were some statistically significant ethnic-group differences that are worth noting based on their effect sizes (i.e., $d \pm 0.40$). For example, Hispanics had significantly fewer years of education than non-Hispanic Whites regardless of diagnosis. Among the normal-cognitive-function-only group, the effect size for this difference was of medium size ($d = -0.45$), as was the case for the dementia-only group, ($d = -0.52$). Additionally, Hispanics had significantly lower scores on the 30-item BNT than non-Hispanic Whites in the normal-cognitive-function-only and dementia-only groups ($d = -0.61$ and -0.42 respectively).

After centering the continuous covariates and key predictor variables on their means, tolerance and VIF values for the logistic regression analyses were not suggestive of problematic multicollinearity. It was predicted that verbal fluency scores would be a stronger predictor of diagnosis in English-speaking non-Hispanic Whites compared to their Hispanic counterparts.

As can be seen in Table 2, analyses revealed that animal fluency had a statistically significant association with diagnosis in both groups such that lower scores were associated with a dementia diagnosis. For the Hispanic patients, the *OR* was 0.81 (95% confidence interval [CI; 0.69, 0.94]; $p < .01$), and for non-Hispanic Whites, this value was 0.83 (95% CI [0.74, 0.93]; $p < .01$). These *ORs* were not statistically significantly different from one another, $z = -0.26, p = .80$, which fails to support the hypothesis that animal fluency would be a significantly stronger predictor of diagnosis in non-Hispanic Whites compared to Hispanics. Regarding vegetable fluency, results showed that scores on this test were significantly related to diagnosis among non-Hispanic Whites, with lower scores being predictive of a dementia diagnosis ($OR = 0.80$; 95% CI [0.69, 0.94]; $p < .01$), but not among Hispanics, $p = .09$. The difference in *ORs* across ethnic groups was nonsignificant, $z = 0.48, p = .63$. This finding suggests that lower vegetable fluency scores were associated with a dementia diagnosis among non-Hispanic Whites only but does not provide evidence for a significant ethnic-group difference regarding the predictive value of this language test on diagnosis.

It was also hypothesized that the predictive value of the 30-item BNT scores on diagnosis would be significantly stronger in English-speaking non-Hispanic Whites compared to their Hispanic counterparts. Analyses revealed that this variable was significantly associated with diagnosis in both Hispanics ($OR = 0.90$; 95% CI [0.81, <1.00]; $p < .05$) and non-Hispanic

Whites ($OR = 0.82$; 95% CI [0.68, 0.99]; $p = .04$), with lower scores being associated with a dementia diagnosis. These OR s did not statistically significantly differ, $z = 0.86$, $p = .39$, failing to corroborate the hypothesis that BNT would be a significantly stronger predictor of diagnosis in non-Hispanic Whites compared to Hispanics.

Results from the logistic regression analysis in which primary language was included as a covariate for English-speaking Hispanics are presented in Table 3.

There were no significant differences in terms of the overall pattern of findings for Hispanics with the inclusion of this variable.

Discussion

The primary aim of this study was to test the hypothesis that neuropsychological language tests of semantic fluency and confrontation naming would be weaker predictors of diagnosis among English-speaking Hispanic versus non-Hispanic White outpatients due to clinicians' likely awareness of these tests' limitations when administered in English among many Hispanics (e.g., Bortnik et al., 2013; Marquez de la Plata et al., 2009). Although this hypothesis was not supported based on tests for significant differences in OR s for each test across ethnic groups, there was a pattern of results suggestive of some ethnic-group differences. Vegetable fluency significantly predicted diagnosis in non-Hispanic Whites only, whereas animal fluency and the 30-item BNT were comparably significant predictors of diagnosis among both ethnic groups. Neuropsychological language tests frequently lack representative normative data and thus often have reduced validity for Hispanics because of many ethnocultural, education, and language factors (e.g., Manly & Espino, 2004). Because the norms for these tests disadvantage Hispanics (as was the case in this study given their nearly consistently significantly lower mean language test scores), it seemed reasonable to hypothesize that the relation of these tests with diagnosis would be significantly weaker for Hispanics. However, taking into account the unique limitations of this study as will be described below, these results did not corroborate this assumption.

Before discussing these results in further detail, it should be emphasized that these findings do not suggest that certain language tests are better predictors of actual diagnosis within and across ethnic groups. Rather, they suggest that certain tests are ultimately associated with the diagnoses made by clinicians in this particular study in the absence of a formal measure of English-language proficiency. Future research should examine the predictive role of these and other language tests in Hispanics using formal measures of English-language abilities, such as literacy and language dominance tests or other tests associated with linguistic dominance using cultural background (e.g., Marin, Sabogal, Marin, Otero-Sabogal, & Perez-Stable, 1987). This suggestion is particularly relevant as self-report tends to be unreliable, especially regarding assessment of skills and abilities (e.g., Merritt, Smith, & Di Renzo, 2005), including second-language skills (e.g., Heilenman, 1990).

In addition to other potentially important diagnostic variables that were not directly assessed in this study, clinicians may have still relied on performance on the 30-item BNT and animal fluency when assessing for dementia among English-speaking Hispanics despite their

known limitations. Several possible reasons may account for this finding. First, neuropsychological tests are relatively more objective measures of cognition compared to other frequently used pieces of diagnostic information such as informant reports, which can be subject to bias. As such, clinicians may have still been inclined to rely on these language tests in the diagnostic process. Second, in the event that clinicians did in fact rely on them, it is possible that they informally adjusted English-speaking Hispanic patients' scores on these tests to account for their limitations as has been shown in prior studies to occur in some cases (e.g., Bortnik et al., 2013; Marquez de la Plata et al., 2009). No firm conclusion can be made about clinicians informally adjusting participants' scores, as no formal inquiry was made regarding administration practices in the current study. Nonetheless, the finding that Hispanic patients' raw mean scores on these two language tests were consistently lower (i.e., 0.91-1.28 points lower for animal fluency and 2.03-3.62 points lower for BNT) than their non-Hispanic White counterparts regardless of diagnosis provides some support for this supposition, as clinicians may have been aware of this systematic performance difference across ethnic groups on some level. Third, the Hispanic patients in this study were proficient enough in English to be evaluated in English versus another language (e.g., Spanish) given that study personnel deemed it appropriate to assess them using the English rather than Spanish Uniform Data Set module. Furthermore, they had a relatively high mean level of education (13.21 years). Therefore, clinicians may have been more inclined to weigh these language tests more heavily, albeit still to a somewhat lesser degree than in non-Hispanic Whites. This hypothesis could not be tested, as this dataset does not contain specific measures related to English-language ability (e.g., number of years speaking English) or education variables (e.g., location of education). It is likely that higher levels of English-language proficiency and education among the Hispanics in this study resulted in better test performance and more similar performance to non-Hispanic Whites, which may not be characteristic of the larger English-speaking Hispanic American population.

In contrast to vegetable fluency, the association of animal fluency with diagnosis did not differ significantly across ethnic groups, which may have in part been driven by the lack of significantly different mean level scores on this measure across ethnic groups among the dementia-only sample (the only case in which Hispanics did not have lower mean scores than non-Hispanic Whites on any of the three tests). Acevedo et al. (2000) found that English speakers with normal cognitive function produced more vegetable responses than their Spanish-speaking counterparts but that the number of animal responses did not differ significantly across these groups. The authors suggested that this difference in verbal fluency across categories may be due to ethnocultural influences on the familiarity, salience, and relevance of items in this category. Specifically, the Hispanics in their study (as well as the English-speaking Hispanics in the current study) may have had access to a smaller variety of vegetables in the US compared to non-Hispanic Whites either due to socioeconomic factors or culturally-influenced food preferences and practices. It is also possible that the English-speaking Hispanics in the current study may have been exposed to some vegetables unique to their nations of origin for which there is no English translation. Furthermore, Hispanics may be less likely to refer to vegetables in English than in Spanish due to culturally-influenced kitchen behaviors and daily norms. As clinicians may gradually become aware of the relatively poorer performance on vegetable versus animal fluency

among Hispanics assessed in English, they may be more hesitant to allow results on this test to influence diagnosis, thus limiting the predictive utility of this measure. Of note, there was a trend toward statistical significance for vegetable fluency's association with diagnosis in Hispanics ($p = .09$), including when covarying primary language ($p = .10$), which may merit closer attention in future similar research.

English-speaking Hispanics who reported having a language other than English (e.g., Spanish) as their primary language were significantly more likely to have been assigned a dementia diagnosis than Hispanics who reported English as their primary language. This finding lends some support to the notion that Hispanics may have performed more poorly on the language tests in this study for this reason, which could thereby increase the chances of clinical misdiagnosis. Nonetheless, in the context of the multivariate regression analysis, primary language was not a significant predictor of diagnosis among Hispanics, suggesting that other factors associated with not having English as a primary language may have been primarily driving this within-ethnic-group diagnostic difference. For example, perhaps primary language was associated with factors such as number of years living in and acculturation to the US that were not assessed in the current study, which could influence scores on language tests and, thus, diagnosis. It may also be likely that bilingualism, which was not measured in this study, may be driving this ethnic-group difference in the role of vegetable fluency on diagnosis. The field would benefit from future studies examining specific factors that may influence the effects of bilingualism and English-language proficiency on diagnosis via language test performance.

The limitations of this study should be noted. First, the external validity of these results may be restricted, as this study's participants represent a convenience sample derived from university-based dementia clinics. Second, although there was sufficient statistical power for the main hypotheses, there was inadequate power to conduct separate analyses based on English-speaking Hispanic patients' primary language and specific Hispanic subgroups, which could result in different findings for this diverse ethnic group. Third, primary language was assessed informally, and this dataset lacked richer information on bilingualism, native language, and English-language proficiency, literacy, and utilization levels other than self-reported primary language. Nonetheless, participants were also eligible to complete the testing using a Spanish module if both they and study personnel believed that this option would be more suitable. Harris, Muñoz, and Llorente (2008) noted that language preference is not synonymous with language proficiency, which is known to impact neuropsychological test performance. Therefore, future research should strive to include these influential variables to examine the role of English-language proficiency on test performance in bilinguals. Fourth, there were only three language tests included in this data set. In addition, the availability of only the 30 odd-numbered items of BNT may have degraded the psychometric properties of this test. Moreover, BNT is known to have limited utility in bilinguals (Roberts et al., 2002; Strauss et al., 2006) even among those proficient in English. Nonetheless, this 30-item version has received empirical support for its reliability and validity (Williams et al., 1989), although not yet among a purely English-speaking Hispanic sample. In addition, BNT is frequently used in assessment by clinicians, and the aim of this study was to examine how these tests are associated with clinicians' diagnoses.

Future research should examine the role of the complete BNT or similar measures that have been validated on and normed with bilingual Hispanics such as the Modified BNT-Spanish (Pontón et al., 1996). Fifth, there was likely some variation in the assessment and diagnostic procedures across sites that could result in reduced reliability, even though the study protocol provided standardized guidance on how to conduct and score assessments and approach diagnosis. Thus, there are probable differences in the training and types of personnel who administered tests and the potential for drift within and across sites. Fortunately, Generalized Estimating Equations statistically accounts for across-site variations in study procedures, test administration, and participant characteristics. Finally, this dataset does not include any variables tapping clinicians' views of the utility of the language tests examined in this study or the extent to which they weighed them across ethnic groups. In spite of these limitations, the current study has several strengths, such as its use of multisite, nationwide data that employed standardized techniques, its rather large and diverse English-speaking Hispanic outpatient sample, and its investigation of the role of multiple language tests in dementia assessment.

In conclusion, this study has important clinical implications regarding differences in the clinical diagnosis of dementia associated with neuropsychological language tests across English-speaking Hispanics and non-Hispanic Whites. First, animal fluency and the 30-item BNT were found to be useful predictive measures in diagnosing both Hispanics and non-Hispanic Whites. Conversely, vegetable fluency was inconsequential in the diagnostic process for Hispanics. As such, animal fluency and BNT may be preferable measures when assessing for dementia in English across these groups (in conjunction with other suitable measures and clinical evaluation) in contrast to vegetable fluency, which may be comparably more affected by ethnocultural and linguistic factors, as previously discussed. Second, clinicians in this study may have been mindful of the reduced validity of language tests among English-speaking Hispanics and therefore informally adjusted BNT and animal fluency scores or gave minimal weight to vegetable fluency compared to non-Hispanic Whites. While well-intentioned, this may be leading to methodical variations in diagnosis within and across ethnic groups. At present, clinicians should aim to employ the most suitable available norms for each patient. The field would benefit from more representative normative data on language tests that account for differences in factors such as English-language proficiency, bilingualism, and number of years living in the US and speaking English to help improve test validity among diverse groups.

Acknowledgments

These data derived from work supported by the National Alzheimer's Coordinating Center Grant [#U01 AG016976]. This content is solely the responsibility of the author and does not necessarily represent the official views of the National Institutes of Health or Aging.

References

- Acevedo A, Loewenstein DA, Barker WW, Harwood DG, Luis C, Bravo M, Duara R. Category Fluency Test: Normative data for English- and Spanish-speaking elderly. *Journal of the International Neuropsychological Society*. 2000; 6:760–769. doi:10.1017/S1355617700677032. [PubMed: 11105466]

- Agresti, A. Modeling correlated, clustered responses. In: Agresti, A., editor. An introduction to categorical data analysis. 2nd ed. John Wiley & Sons, Inc; Hoboken, NJ: 2007. p. 276-296.
- Allison, PD. Binary logistic regression: Details and options. In: Allison, PD., editor. Logistic regression using SAS: Theory and application. 2nd ed. SAS Institute Inc; Cary, NC: 2012. p. 39-108.
- Altman DG, Bland JM. Interaction revisited: The difference between two estimates. *British Medical Journal*. 2003; 326:219. doi:10.1136/bmj.326.7382.219. [PubMed: 12543843]
- American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Author; Washington: 2000. p. 147-154. text revision
- Alzheimer's Association. Hispanics/Latinos and Alzheimer's disease. Author; Chicago: 2004.
- Bortnik KE, Boone KB, Wen J, Lu P, Mitrushina M, Razani J, Maury T. Survey results regarding use of the Boston Naming Test: Houston, we have a problem. *Journal of Clinical and Experimental Neuropsychology*. 2013; 35:857-866. doi:10.1080/13803395.2013.826182. [PubMed: 23998683]
- Burke DM, MacKay DG, Worthley JS, Wade E. On the tip of the tongue: What causes word finding failures in young and older adults? *Journal of Memory and Language*. 1991; 30:542-579. doi: 10.1016/0749-596X(91)90026-G.
- Folstein MF, Folstein SE, McHugh PR. "Mini-Mental State." A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*. 1975; 12:189-198. doi: 10.1016/0022-3956(75)90026-6. [PubMed: 1202204]
- Gollan TH, Montoya RI, Werner GA. Semantic and letter fluency in Spanish-English bilinguals. *Neuropsychology*. 2002; 16:562-576. doi:10.1037//0894-4105.16.4.562. [PubMed: 12382994]
- Gollan TH, Silverberg NB. Tip-of-the-tongue states in Hebrew-English bilinguals. *Bilingualism: Language and Cognition*. 2001; 4:63-83. doi:10.1017/S136672890100013X.
- Harris, JG.; Muñoz, MR.; Llorente, AM. Considerations in using the WISC-IV with Hispanic children. In: Prifitera, A.; Saklofske, DH.; Weiss, LG., editors. WISC-IV clinical assessment and intervention. 2nd ed. Academic Press; San Diego, CA: 2008. p. 410-444.
- Heilenman LK. Self-assessment of second language ability: The role of response effects. *Language Testing*. 1990; 7:174-201. doi:10.1177/026553229000700204.
- Hernandez AE, Kohnert KJ. Aging and language switching in bilinguals. *Aging, Neuropsychology, and Cognition*. 1999; 6:69-83. doi:10.1076/anec.6.2.69.783.
- Ivanova I, Costa A. Does bilingualism hamper lexical access in speech production? *Acta Psychologica*. 2008; 127:277-288. doi:10.1016/j.actpsy.2007.06.003. [PubMed: 17662226]
- Kaplan, E.; Goodglass, H.; Weintraub, S. The Boston Naming Test. Lea and Febiger; Philadelphia: 1983.
- Kohnert KJ, Hernandez AE, Bates E. Bilingual performance on the Boston Naming Test: Preliminary norms in Spanish and English. *Brain and Language*. 1998; 65:422-440. doi:10.1006/brln.1998.2001. [PubMed: 9843612]
- Le Carret N, Lafont S, Letenneur L, Dartigues J-F, Mayo W, Fabrigoule C. The effect of education on cognitive performances and its implication for the constitution of the cognitive reserve. *Developmental Neuropsychology*. 2003; 23:317-337. doi:10.1207/S15326942DN2303_1. [PubMed: 12740188]
- Manly JJ, Byrd D, Touradji P, Sanchez D, Stern Y. Literacy and cognitive change among ethnically diverse elders. *International Journal of Psychology*. 2004; 39:47-60. doi: 10.1080/00207590344000286.
- Manly JJ, Espino DV. Cultural influences on dementia recognition and management. *Clinics in Geriatric Medicine*. 2004; 20:93-119. doi:10.1016/j.cger.2003.10.004. [PubMed: 15062490]
- Marin G, Sabogal F, Marin BV, Otero-Sabogal R, Perez-Stable EJ. Development of a short acculturation scale for Hispanics. *Hispanic Journal of Behavioral Sciences*. 1987; 9:183-205. doi: 10.1177/07399863870092005.
- Marquez de la Plata C, Arango-Lasprilla JC, Alegret M, Moreno A, Tárraga L, Lara M, Cullum CM. Item analysis of three Spanish naming tests: A cross-cultural investigation. *NeuroRehabilitation*. 2009; 24:75-85. doi:10.3233/NRE-2009-0456. [PubMed: 19208960]
- McKhann G, Drachman D, Folstein M, Katzman R, Price D, Stadlan EM. Clinical diagnosis of Alzheimer's disease: Report of the NINCDS-ADRDA Work Group under the auspices of

- Department of Health and Human Services Task Force on Alzheimer's Disease. *Neurology*. 1984; 34:939–944. doi:10.1212/WNL.34.7.939. [PubMed: 6610841]
- Merritt K, Smith KD, Di Renzo JC Jr. An investigation of self-reported computer literacy: Is it reliable? *Issues in Information Systems*. 2005; 6:289–295. Retrieved from: http://iacis.org/iis/2005/Merritt_Smith_DiRenzo.pdf.
- Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of functional activities in older adults in the community. *Journal of Gerontology*. 1982; 37:323–329. doi:10.1093/geronj/37.3.323. [PubMed: 7069156]
- Pontón MO, Satz P, Herrera L, Ortiz F, Urrutia CP, Young R, Namerow N. Normative data stratified by age and education for the Neuropsychological Screening Battery for Hispanics (NeSBHIS): Initial report. *Journal of the International Neuropsychological Society*. 1996; 2:96–104. (0. doi: 10.1017/S1355617700000941. [PubMed: 9375194]
- Portocarrero JS, Burright RG, Donovick PJ. Vocabulary and verbal fluency of bilingual and monolingual college students. *Archives of Clinical Neuropsychology*. 2007; 22:415–422. doi: 10.1016/j.acn.2007.01.015. [PubMed: 17336036]
- Reitan RM. Validity of the Trail Making test as an indicator of organic brain damage. *Perceptual and Motor Skills*. 1958; 8:271–276. doi:10.2466/pms.1958.8.3.271.
- Roberts PM, Garcia LJ, Desrochers A, Hernandez D. English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology*. 2002; 16:635–645. doi: 10.1080/02687030244000220.
- Román GC, Tatemichi TK, Erkinjuntti T, Cummings JL, Masdeu JC, Garcia JH, Scheinberg P. Vascular dementia: Diagnostic criteria for research studies: Report of the NINDS-AIREN International Workshop. *Neurology*. 1993; 43:250–260. doi:10.1212/WNL.43.2.250. [PubMed: 8094895]
- Rosselli M, Ardila A, Araujo K, Weekes VA, Caracciolo V, Padilla M, Ostrosky-Solís F. Verbal fluency and repetition skills in healthy older Spanish-English bilinguals. *Applied neuropsychology*. 2000; 7:17–24. doi:10.1207/S15324826AN0701_3. [PubMed: 10800624]
- Sayegh P, Knight BG. Assessment and diagnosis of dementia in Hispanic and non-Hispanic White outpatients. *The Gerontologist*. 2013a Advance online publication. doi:10.1093/geront/gns190.
- Sayegh P, Knight BG. Functional Assessment and Neuropsychiatric Inventory Questionnaires: Measurement invariance across Hispanics and non-Hispanic Whites. *The Gerontologist*. 2013b Advance online publication. doi:10.1093/geront/gnt026.
- Strauss, E.; Sherman, EMS.; Spreen, O. A compendium of neuropsychological tests: Administration, norms, and commentary. 3rd ed. Oxford University Press; New York: 2006. p. 891-962.
- Suárez-Orozco, MM.; Pérez, MM. Introduction. In: Suárez-Orozco, MM.; Pérez, MM., editors. *Latinos: Remaking America*. University of California Press; Berkeley: 2002. p. 1-43.
- Valle R, Lee B. Research priorities in the evolving demographic landscape of Alzheimer's Disease and associated dementias. *Alzheimer Disease and Associated Disorders*. 2002; 16:S64–S76. doi: 10.1097/00002093-200200002-00006. [PubMed: 12351917]
- Wechsler, D. Wechsler Memory Scale—Revised Manual. Psychological Corporation; San Antonio: 1987a.
- Wechsler, D. Wechsler Adult Intelligence Scale—Revised. Psychological Corporation; San Antonio: 1987b.
- Weintraub S, Salmon D, Mercaldo N, Ferris S, Graff-Radford NR, Chui H, Morris JC. The Alzheimer's disease centers' uniform data set (Uniform Data Set): The neuropsychological test battery. *Alzheimer Disease and Associated Disorders*. 2009; 23:91–101. doi:10.1097/WAD.0b013e318191c7dd. [PubMed: 19474567]
- Williams BW, Mack W, Henderson VW. Boston Naming Test in Alzheimer's disease. *Neuropsychologia*. 1989; 27:1073–1079. doi:10.1016/0028-3932(89)90186-3. [PubMed: 2797414]

Table 1
Descriptive Statistics of and Group Differences in Demographic and Key Variables, by Ethnicity and Diagnosis

	All (<i>N</i> = 872), <i>n</i> (%)	NCF (<i>n</i> = 400), <i>n</i> (%)	Dementia (<i>n</i> = 472), <i>n</i> (%)	NCF v. Dementia, <i>p</i> value
Sex (Women)				
Hispanic (<i>N</i> = 436)	259 (59.40)	138 (69.35)	121 (51.05)	<.001
NHW (<i>N</i> = 436)	248 (56.88)	128 (63.68)	120 (51.06)	.008
Prob (<i>H</i> ₀)	.450	.230	.998	---
Primary lang. (English)				
Hispanic (<i>N</i> = 436)	297 (68.12)	155 (77.89)	142 (59.29)	<.001
NHW (<i>N</i> = 436)	430 (98.62)	198 (98.51)	232 (98.72)	.847
Prob (<i>H</i> ₀)	<.001	<.001	<.001	---
Age (years)				
Hispanic (<i>N</i> = 436)	71.67 (10.28)	69.61 (9.64)	73.40 (10.49)	<.001
NHW (<i>N</i> = 436)	72.28 (10.74)	71.94 (10.58)	72.57 (10.90)	.541
Prob (<i>H</i> ₀)	.395	.022	.399	---
Education (years)				
Hispanic (<i>n</i> = 433)	13.21 (3.93)	13.96 (3.63)	12.57 (4.07)	<.001
NHW (<i>n</i> = 430)	14.93 (3.18)	15.44 (2.97)	14.50 (3.29)	.002
Prob (<i>H</i> ₀)	<.001	<.001	<.001	---
Total FAQ score				
Hispanic (<i>N</i> = 436)	9.79 (10.95)	0.36 (1.12)	17.70 (9.06)	<.001
NHW (<i>N</i> = 436)	8.73 (9.97)	0.61 (2.02)	15.68 (8.73)	<.001
Prob (<i>H</i> ₀)	.138	.126	.014	---
Animal fluency				
Hispanic (<i>n</i> = 410)	13.83 (6.87)	18.76 (5.03)	9.26 (4.90)	<.001
NHW (<i>n</i> = 409)	15.00 (7.32)	20.04 (5.89)	10.17 (4.91)	<.001
Prob (<i>H</i> ₀)	.018	.020	.058	---
Vegetable fluency				
Hispanic (<i>n</i> = 424)	8.99 (5.66)	13.39 (4.10)	5.17 (3.72)	<.001

	All (<i>N</i> = 872), <i>n</i> (%)	NCF (<i>n</i> = 400), <i>n</i> (%)	Dementia (<i>n</i> = 472), <i>n</i> (%)	NCF v. Dementia, <i>p</i> value
NHW (<i>n</i> = 424)	10.07 (5.78)	14.38 (4.31)	6.29 (3.98)	<.001
Prob (<i>H</i> ₀)	.006	.020	.002	---
Boston Naming Test *				
Hispanic (<i>n</i> = 420)	19.78 (8.39)	25.49 (4.03)	14.73 (7.99)	<.001
NHW (<i>n</i> = 424)	22.58 (8.30)	27.52 (2.50)	18.35 (9.17)	<.001
Prob (<i>H</i> ₀)	<.001	<.001	<.001	---

Notes. NCF = normal cognitive function; NHW = non-Hispanic White; FAQ = Functional Assessment Questionnaire; Prob (*H*₀) = null hypothesis that group differences are not significant.

* 30 odd-numbered items.

Table 2

Logistic Regression of Neuropsychological Language Test Performance on Diagnosis with Covariates, by Ethnic Group

Hispanics	OR	SE	Wald chi-square	p value	95% CI
Age (years)	0.975	0.032	0.640	0.425	0.915-1.038
Female sex	0.365	0.621	2.624	0.105	0.108-1.234
Education (years)	1.022	0.086	0.068	0.799	0.864-1.209
Total FAQ score	2.423	0.164	29.160	<0.001	1.757-3.340
Animal fluency	0.806	0.079	7.453	0.006	0.690-0.941
Vegetable fluency	0.850	0.095	2.924	0.087	0.705-1.024
Boston Naming Test *	0.899	0.054	3.881	0.049	0.808-0.999
Non-Hispanic Whites	OR	SE	Wald chi-square	p value	95% CI
Age (years)	0.937	0.038	3.028	0.082	0.870-1.008
Female sex	0.546	0.587	1.061	0.302	0.173-1.724
Education (years)	1.044	0.088	0.240	0.627	0.878-1.241
Total FAQ score	1.691	0.124	17.893	<0.001	1.325-2.156
Animal fluency	0.827	0.060	10.112	0.001	0.735-0.929
Vegetable fluency	0.801	0.079	7.896	0.005	0.686-0.935
Boston Naming Test *	0.817	0.097	4.368	0.037	0.676-0.988

Notes. FAQ = Functional Assessment Questionnaire; OR = odds ratio; SE = standard error; 95% CI = 95% confidence interval.

* 30 odd-numbered items.

Table 3

Logistic Regression of Neuropsychological Language Test Performance on Diagnosis with Covariates Including Primary Language for Hispanics

Hispanics	<i>OR</i>	<i>SE</i>	Wald chi-square	<i>p</i> value	95% CI
Age (years)	0.971	0.032	0.846	0.359	0.913-1.034
Female sex	0.336	0.621	3.098	0.079	0.099-1.132
Education (years)	1.023	0.088	0.063	0.799	0.861-1.215
Primary language (English)	1.829	0.371	2.657	0.104	0.884-3.784
Total FAQ score	2.496	0.159	33.178	<0.001	1.829-3.407
Animal fluency	0.787	0.089	7.344	0.007	0.661-0.936
Vegetable fluency	0.854	0.096	2.690	0.100	0.708-1.031
Boston Naming Test *	0.890	0.059	3.920	0.048	0.793-0.999

Notes. FAQ = Functional Assessment Questionnaire; *OR* = odds ratio; *SE* = standard error; 95% CI = 95% confidence interval.

* 30 odd-numbered items.