

Cross-Cultural Neuropsychological Assessment: A Comparison of Randomly Selected, Demographically Matched Cohorts of English- and Spanish-Speaking Older Adults*

Diane M. Jacobs, Mary Sano, Steven Albert, Peter Schofield, George Dooneief, and Yaakov Stern
Department of Neurology and Gertrude H. Sergievsky Center, Columbia University,
College of Physicians and Surgeons, New York

ABSTRACT

As the US population of elderly Hispanics continues to grow, there is an increasingly greater need for neuropsychological measures that are appropriate for assessing Spanish-speaking elders. The current study compared the performance of randomly selected, community-based samples of English- and Spanish-speaking elders on a brief neuropsychological test battery. Subject groups were matched for age and education. Multivariate analysis indicated significant group differences on the test battery. English and Spanish speakers scored comparably on many language-based tasks, but Spanish speakers scored significantly lower on almost all of the nonverbal measures. Significant group differences were observed on multiple-choice matching and recognition memory for stimuli from the Benton Visual Retention Test, as well as on Identities and Oddities from the Mattis Dementia Rating Scale, category fluency, and Complex Ideational Material from the Boston Diagnostic Aphasia Examination (BDAE). Results suggest that caution is warranted when using nonverbal as well as verbal measures to assess non-English-speaking individuals.

Neuropsychological evaluation is an important tool in the clinical detection and diagnosis of Alzheimer's disease and other age-associated cognitive disorders. A severe limitation of many widely used neuropsychological instruments, however, is that they are available and standardized only in English. As a result, it can be difficult to assess cognitive impairment in individuals who are not fluent English speakers.

Hispanics are one of the most rapidly growing ethnic subgroups in the United States (Campbell, 1994). Projected population trends indicate that Hispanics will account for one third of the growth in the nation's population between

1993 and 2000, and 38% of the growth between 1993 and 2020 (Campbell, 1994). By the year 2000, Hispanics are expected to be the largest ethnic minority in the US (Davis, Haub, & Willett, 1983). This, coupled with the fact that the life expectancy of persons living in the US, including Hispanics, has continued to improve, has led to a growing increase in the population of Spanish-speaking elders.

To assess cognition in Spanish-speaking elders, clinicians often use translations of English measures, or nonverbal tasks, such as the Performance subtests from the Wechsler Adult Intelligence Scale-Revised (WAIS-R) and tests of

* This study was supported by Federal grants AG07232, AG08702, RR00645, and the Charles S. Robertson Gift for Alzheimer's Disease. The authors are grateful to Dr. Luisa A. Contreras for her helpful comments on an earlier version of this article.

Results of this research were presented in part at the annual meeting of the International Neuropsychological Society, Chicago, IL, February 1996.

Address correspondence to: Diane M. Jacobs, Ph.D., Gertrude H. Sergievsky Center, 630 West 168 Street, New York, NY 10032-3702, USA.

Accepted for publication: November 1, 1996.

figural or facial memory. These practices, however, do not necessarily result in valid assessments of cognition. Previous studies have suggested that tests translated from English and administered in Spanish often are not equivalent to the original English version. For example, several investigations have compared Spanish and English versions of brief mental status screening instruments, such as the Mini-Mental State Examination (Bird, Canino, Stipek, & Shrout, 1987; Valle et al., 1991). Results indicated that Spanish-speaking subjects were more likely than English speakers to miss certain items regardless of their general cognitive functioning. As a result, Spanish-speaking persons, particularly those with limited educational backgrounds, were more likely to be classified as impaired, despite being cognitively intact on clinical examination. Nonverbal neuropsychological measures also may underestimate abilities of elderly Hispanics. Lopez and Taussig (1991) found that cognitively nonimpaired Spanish-speaking elders scored significantly lower than their English-speaking peers on the Block Design subtest of the WAIS-R. Loewenstein and colleagues (Loewenstein, Arguelles, Barker, & Duara, 1993) reported that moderately demented Spanish speakers with Alzheimer's disease (AD) scored lower than moderately demented English speakers with AD on the Block Design and Object Assembly subtests of the WAIS-R, although mildly demented Spanish and English speakers did not differ significantly on these measures.

Methodological problems have limited somewhat the conclusions that can be drawn from many previous comparisons of English and Spanish speakers on neuropsychological tests. First, the generalizability of results often is limited because subjects are from convenience samples, such as self-referred patients or research volunteers. For generalizability and wider applicability, representative samples of subjects are required. Further, demographic differences between Spanish and English speakers often are confounded with the overall group effect. In particular, Spanish speakers, as a group, often have lower levels of educational attainment than do English speakers, making interpretation of group differences difficult. Analysis of covariance typ-

ically has been used to statistically control for such demographic differences between groups; however, this technique was designed to correct for errors of randomization, not to statistically equate groups on a covariate. As a result, analysis of covariance may fail to adequately control for preexisting differences between groups.

The goal of this study was to compare the performance of a demographically matched, randomly selected, community-based sample of nondemented English- and Spanish-speaking elders on a variety of neuropsychological measures in an effort to determine which measures yielded comparable performance in English and Spanish speakers and which did not. Results presented here were derived from secondary analyses of a neuropsychological test battery that was administered during the course of a large-scale, community-based epidemiological study of dementia.

METHODS

Subjects

All subjects were participants in the Washington Heights-Inwood Columbia Aging Project (WHICAP), a community-based, epidemiological study of dementia in Northern Manhattan, NY. The community from which subjects were drawn is ethnically diverse, and approximately 40% of residents are primarily or exclusively Spanish-speaking. Almost all businesses and institutions in the neighborhoods of Washington Heights and Inwood are bilingual or Spanish-speaking; thus, many of the Hispanic elders in our cohort speak little or no English. Hispanic subjects vary in country of origin, but are predominantly from the Dominican Republic, Cuba, and Puerto Rico.

The WHICAP sample is a 50% random sample of elders residing in Washington Heights and Inwood: the Health Care Finance Association provided a 50% sample of names and addresses of all Medicare recipients in the three ZIP codes in North Manhattan. All records for persons residing out of the area and persons with birth dates making them less than 65 years of age were dropped. Overall, WHICAP has above a 70% participation rate. Participation is slightly higher in Hispanic subjects.

Only WHICAP subjects who showed no signs of dementia or cognitive impairment were included in the current analyses. This determination was

made on the basis of a neurologist's clinical examination which included mental status testing and a rating of day-to-day function, but without knowledge of the subject's performance on neuropsychological measures. Subjects with major depression or history of Parkinson's disease, stroke, head injury with loss of consciousness, or alcohol abuse were excluded. All included subjects had complete data on the neuropsychological test battery.

Stratified random sampling was used to create two subject groups that were comparable in terms of distribution of age and education. Subjects were classified as English- or Spanish-speaking based upon the language in which they chose to complete the neuropsychological examination. WHICAP participants who met inclusion/exclusion criteria were stratified by language (Spanish, English), age (65–70, 71–75, 76–80, 81–85, 86–90), and years of education (0–3, 4–7, 8–11, 12–15, 16–20). Within each language by age by education stratum, a random sample of subjects was selected using the Statistical Program for the Social Sciences (SPSS, 1993) sample function. The same number of English and Spanish speakers was selected within each age by education stratum; thereby yielding two groups (one English- and one Spanish-speaking) that had a similar distribution of age and education.

Two-hundred and thirty-six subjects (118 English and 118 Spanish) were selected via the sampling procedure described above. Demographic characteristics of English and Spanish speakers are presented in Table 1. English and Spanish groups were well-matched for age, years of education, and male/female distribution.

Ninety-eight percent of the subjects in the Spanish-speaking group were immigrants to the United States, primarily from the Dominican Republic (46%), Cuba (27%), and Puerto Rico (17%). The majority of these subjects have resided in the United States for more than 15 years, with only 1 subject immigrating after 1980. When asked how well they speak English, 43% of Spanish speakers responded "not at all", 33% "not well", 11% "well", and 13% "very well". Ninety-seven per-

cent of the subjects in the Spanish-speaking group identified Spanish as the predominant language spoken at home.

Seventy-seven percent of the English-speaking subjects were born and raised in the United States. The majority of the English-speaking immigrants were from European countries, and all of them immigrated prior to 1980. Ninety-eight percent of English-speaking subjects spoke English well or very well, and 89% spoke predominantly English at home.

Procedures

By virtue of their participation in WHICAP, all subjects received a standardized neurological and neuropsychological evaluation. A behavioral neurologist conducted a physical and neurological examination and recorded each subject's medical history. All available ancillary information, including medical records and imaging studies, was used in the evaluation. In addition, a brief (approx. 1 hr) neuropsychological test battery was administered. Physical and neuropsychological examinations generally were performed in the subject's home; however, if the subject preferred to be evaluated at the medical center, travel expenses were reimbursed.

The neuropsychological battery was designed to assess a range of cognitive functions that typically are affected in dementia. It consists primarily of subtests and short forms of widely used neuropsychological measures. Specific cognitive functions assessed and neuropsychological tests administered include the following:

Word list learning and memory

The Selective Reminding Test (SRT; Buschke & Fuld, 1974) was administered. Subjects were given six trials to learn a list of 12 unrelated words. After each recall attempt, subjects were reminded only of those words that had not been successfully recalled. To assess long-term retention of the word list, 15-min delayed free recall was assessed, followed by a multiple-choice recognition task.

Table 1. Demographic Characteristics (Means and Standard Deviations).

	English (<i>n</i> = 118)		Spanish (<i>n</i> = 118)		Statistic; Signif.
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
Age (years)	75.07	5.90	74.91	5.71	<i>t</i> = .21; <i>p</i> = .83
Education (years)	8.85	3.78	8.41	3.98	<i>t</i> = .87; <i>p</i> = .38
Sex (% Female)	75		72		χ^2 = .09; <i>p</i> = .77

Nonverbal memory. A multiple-choice version of the Benton Visual Retention Test (BVRT; Benton, 1955) was used to assess nonverbal memory. Subjects viewed a geometric design for 10 s. It was then removed from view, and the subject was asked to recognize the design in a four-choice multiple-choice array. Stimuli corresponded to Form D of the original Benton Visual Retention Test.

Orientation. The 10 orientation items from the Mini-Mental State Examination (Folstein, Folstein, & McHugh, 1975) were used to assess orientation to time and place.

Verbal reasoning. The Similarities subtest of the Wechsler Adult Intelligence Scale-Revised (Wechsler, 1981), which requires subjects to identify relevant similarities or superordinate categories for paired items, was administered.

Nonverbal reasoning. The Identities and Oddities subtest of the Mattis Dementia Rating Scale (Mattis, 1976) was used to assess nonverbal reasoning. In this test, subjects identify which two of three visually presented stimuli are the same. After all eight triads are completed, the same items are administered again, and the subject is required to identify the one item that is different.

Naming. A 15-item version of the Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983), a test of visual confrontation naming, was used to assess word-finding ability.

Letter fluency. We administered the Controlled Word Association Test from the Multilingual Aphasia Examination (Benton & Hamsher, 1976). Subjects were instructed to generate as many words as possible in 1 min that begin with a given letter of the alphabet, excluding proper nouns and the same word with a different suffix. Three letters of progressively increasing associative difficulty, as defined by the relative frequency of words beginning with the letter stimuli, were presented as stimuli. English-speaking subjects generated words beginning with the letters *C*, *F*, and *L*, whereas Spanish-speaking subjects generated words beginning with the letters *P*, *S*, and *V*. Different letters were used for Spanish- and English-speaking subjects to control for word-frequency differences across the two languages. Relative frequencies of words beginning with each letter were determined from standard dictionaries for each language as reported in the *Manual for the Multilingual Aphasia Examination* (Benton & Hamsher, 1976). The letters *C/F/L* in English and *P/B/V* in Spanish are comparable in terms of mean frequency rank; however, because *B* and *V* are allophonic in Spanish, we substituted *S* (the letter with moderate associative difficulty from Form II of the Spanish Controlled Word Association Test)

for *B* in the Spanish task. As a result, the mean associative difficulty was very slightly less difficult on the Spanish version of this task.

Category fluency. All subjects generated exemplars in the categories animals, foods, and clothing; 60 s was allowed for each category. For animal fluency, we used the instructions from the Boston Diagnostic Aphasia Examination (BDAE) (Goodglass & Kaplan, 1983); subjects were instructed to name any type of animal, but were provided superordinate categories (animals from the farm, jungle, ocean, or house pets) to assist retrieval. For clothing and food fluency, subjects were instructed to name anything one could wear ("qualquier cosa que se pueda usar para vestir") or eat ("qualquier cosa que se pueda comer"), respectively.

Auditory comprehension. The first six items of the Complex Ideational Material subtest of the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1983) were used to assess comprehension of spoken language.

Repetition. Subjects were asked to repeat the high-frequency phrases from the Boston Diagnostic Aphasia Examination Repetition of Phrases subtest (Goodglass & Kaplan, 1983).

Visuoconstructional skills. Subjects copied five designs from the Rosen Drawing Test (Rosen, 1981) ranging in difficulty from simple geometric shapes to overlapping, parallel, and three-dimensional figures.

Visuoperceptual skills. Subjects matched a target design to the same design presented simultaneously in a four-choice multiple-choice array containing the target along with three distractors. Target stimuli corresponded to Form C of the original Benton Visual Retention Test (BVRT; Benton, 1955).

All interview questions, test instructions, and stimuli were translated into Spanish by a committee of native Spanish speakers from Cuba, Puerto Rico, Spain, and the Dominican Republic, and then back-translated to ensure accuracy. Test items were translated literally. Where necessary, scoring criteria were modified so as to give credit for responses reflecting regional idioms. For example, on the Boston Naming Test in addition to "silla de ruedas" (chair of wheels), "silla de invalidos" (chair of invalids) was an acceptable response to the stimulus depicting a wheelchair; similarly, in addition to "globo" (globe), "bola del mundo" (ball of the world) and "mapamundi" (world map) were acceptable responses for globe. Homophones on the Controlled Word Association Test (e.g., "sera" vs. "cera") were counted as correct (i.e., the examiner assumed that the subject was provid-

ing the word with the task appropriate initial letter).

Evaluations were conducted in either English or Spanish, based on the subject's opinion of which language would yield better performance. Examiners were balanced bilinguals, who spoke both English and Spanish daily with friends, family, and colleagues. Most examiners were native Spanish speakers or first-generation American citizens who were raised in Spanish-speaking households. All examiners who were not native Spanish speakers were interviewed in Spanish by native speaking members of the research team to ensure language proficiency.

Data Analysis

Multivariate analysis of variance (MANOVA) was used to compare the overall neuropsychological test performance of English and Spanish speakers. Because of the stratified random sampling technique, there was no need to include age and education as covariates; however, we were interested in examining possible interactive effects of language with age or education. Dichotomous variables for age (< 75 years, > 75 years) and education (< 9 years, > 9 years) were created based upon the median values for the entire cohort. The multivariate model examined the overall effect of language and the interactive effects of language by age and language by education. Fourteen measures from the test battery were included as dependent variables in a two-group MANOVA as follows: total raw scores for immediate recall, delayed recall, and delayed recognition memory from the SRT; BVRT matching and recognition memory; 15-item Boston Naming; abbreviated BDAE Repetition; abbreviated BDAE Complex Ideational Material; Mattis Identities and Oddities; MMSE Orientation; and number correct on Rosen Drawing; age-corrected scaled score on WAIS-R Similarities; and mean number of words generated over three 60-s trials for both letter and category fluency. Post hoc univariate ANOVAs were examined to determine which neuropsychological measures contributed significantly to the overall multivariate effect. A Bonferroni corrected alpha level of $p < .004$ (i.e., $.05/14$) was used to determine statistically significant univariate differences.

For those measures yielding significant group differences, we examined the influence of various demographic factors within the Spanish-speaking group on test performance using linear regression analyses. Independent variables for the regression analyses were age; education; English proficiency (English spoken well or very well vs. not well or not at all); year of immigration to the United States

(before 1960 vs. in or after 1960); and country of origin (dummy coded as three variables: Puerto Rican vs. not Puerto Rican, Cuban vs. not Cuban, Dominican vs. not Dominican). Age and education were treated as continuous variables; data on English proficiency and year of immigration were available only as categorical variables. There were no significant differences between Spanish-speaking men and women in terms of age, education, English proficiency, country of origin, or year of immigration; therefore, gender was not included as a factor in these analyses. A Bonferroni corrected alpha level of $p < .007$ (i.e., $.05/7$) was used to determine which variables contributed significantly to the regression equations.

Secondary analyses

We were interested in examining the impact of level of acculturation (i.e., the level to which individuals participate in the values, language, and practices of their native ethnic culture versus those of the dominant culture (Padilla, 1980) on test performance. Unfortunately, no formal measure of acculturation was available in our cohort. Acculturation is strongly associated with language use (Marin & VanOss Marin, 1991); therefore, as an estimate of level of acculturation, we examined further the effect of English fluency on test performance. Subjects who were tested in Spanish but reported that they spoke English "very well" or "well" were pairwise matched for age and education to subjects tested in Spanish who reported that they spoke English "not well" or "not at all" and to subjects tested in English. We compared the performance of these three matched groups on each of the neuropsychological measures using one-way analyses of variance with Bonferroni corrected post-hoc comparisons. Because these analyses were exploratory and power was limited by relatively small sample sizes, we chose not to use an alpha-correction procedure for the overall F tests.

RESULTS

Multivariate analysis of variance comparing English- and Spanish-speaking elders on all of the measures in the test battery yielded a significant effect of language ($F(14, 219) = 5.47$; $p < .0001$), but no interactions of language with age ($F(14, 219) = .81$; $p = .66$) or education ($F(14, 219) = .92$; $p = .54$). Average test scores for English and Spanish speakers on each of the

neuropsychological measures and results of univariate analyses of variance are shown in Table 2. English- and Spanish-speaking elders scored comparably on the majority of our verbally mediated neuropsychological measures, including the Selective Reminding Test, Boston Naming Test, and WAIS-R Similarities. Spanish speakers, however, did score significantly lower than English speakers on category fluency and abbreviated Complex Ideational Material from the BDAE. Spanish speakers also scored significantly lower than English speakers on three of the four nonverbal measures, including Identities and Oddities from the Mattis Dementia Rating Scale, BVRT matching, and BVRT recognition memory.

Linear regression analyses, using test scores for Spanish speakers on each of the five measures which yielded significant univariate group differences as dependent variables (i.e., Mattis Identities and Oddities, Benton Matching and Recognition Memory, Category Fluency, and abbreviated BDAE Complex Ideational Material), produced the following results. Proficiency in English and year of immigration failed to contribute significantly to any of the regression equations. Country of origin contributed significantly to only one regression equation: having immigrated from Cuba was associated with better performance on Benton Matching ($B =$

1.79; $p < .004$). Age was a significant predictor of performance on Mattis Identities and Oddities ($B = -0.11$; $p < .002$) and Benton Recognition Memory ($B = -0.09$; $p < .04$), whereas education was a significant predictor of performance on Benton Matching ($B = 0.17$; $p < .0005$) and abbreviated BDAE Complex Ideational Material ($B = 0.09$; $p < .004$).

The 28 subjects who were tested in Spanish but reported that they spoke English "well" or "very well" were pairwise matched to 28 subjects tested in Spanish who reported that they spoke English "not well" or "not at all," and to 28 subjects tested in English. The resulting group means for age and education were 74 and 10 years, respectively. The distribution of men and women also was comparable across groups. One-way analyses of variance comparing the three matched groups on each of the neuropsychological measures yielded significant group differences for BVRT matching ($F(2, 81) = 3.31$; $p < .05$), BVRT recognition memory ($F(2,81) = 6.88$; $p < .002$), category fluency ($F(2, 81) = 5.25$; $p < .01$), and abbreviated Complex Ideational Material from the BDAE ($F(2, 81) = 7.55$; $p < .001$). Post hoc comparisons revealed that on each of these measures subjects tested in English scored significantly higher than did subjects tested in Spanish who reported that they did not speak English well; subjects tested

Table 2. Mean (*SD*) Neuropsychological Test Scores for English- and Spanish-speaking Elders.

Test (max. score)	English ($n = 118$)	Spanish ($n = 118$)	Univariate <i>F</i> test ($df = 1,232$)
SRT Immediate Recall (72)	36.82 (9.76)	36.70 (9.25)	$F = .10$; $p = .76$
SRT Delayed Recall (12)	5.64 (2.63)	5.17 (2.15)	$F = 2.45$; $p = .12$
SRT Delayed Recog. (12)	10.92 (1.53)	11.03 (1.44)	$F = .31$; $p = .58$
BVRT Recog. Memory (10)	6.79 (1.95)	5.74 (1.99)	$F = 16.82$; $p < .0001$
BVRT Matching (10)	8.30 (1.60)	7.58 (1.93)	$F = 10.64$; $p = .001$
Rosen Drawing (5)	2.73 (.94)	2.60 (1.04)	$F = 1.13$; $p = .29$
Boston Naming (15)	13.42 (1.58)	13.51 (1.60)	$F = .15$; $p = .70$
Letter Fluency	8.43 (4.04)	7.72 (2.95)	$F = 2.80$; $p = .10$
Category Fluency	14.05 (3.81)	12.14 (3.47)	$F = 16.73$; $p < .0001$
BDAE Repetition (8)	7.56 (.77)	7.53 (.85)	$F = .07$; $p = .79$
BDAE Complex Ideation (6)	5.25 (1.08)	4.60 (1.23)	$F = 19.35$; $p < .0001$
WAIS-R Similarities (19)	7.72 (2.93)	7.61 (2.47)	$F = .20$; $p = .65$
Identities & Oddities (16)	14.25 (1.95)	13.47 (2.14)	$F = 9.28$; $p = .003$
MMSE Orientation (10)	9.42 (1.06)	9.44 (1.02)	$F = .05$; $p = .82$

Note. SRT = Selective Reminding Test; BVRT = Benton Visual Retention Test; BDAE = Boston Diagnostic Aphasia Examination; WAIS-R = Wechsler Adult Intelligence Scale - Revised; MMSE = Mini-Mental State Examination.

in Spanish who reported that they spoke English "well" or "very well" did not differ significantly from either of the other two groups.

DISCUSSION

Subject characteristics (e.g., age, education, language, culture) can dramatically affect performance on neuropsychological tests; therefore, it is necessary to take into account the influence of these factors when interpreting test results. The neuropsychological effects of age and education have been well documented (e.g., Heaton, Grant, & Matthews, 1986; Zarit, Miller, & Kahn, 1978). We compared the test performance of English- and Spanish-speaking elders after matching groups for age and education. Spanish speakers from our demographically matched, randomly sampled, community-based cohort scored significantly lower on a neuropsychological test battery than did their English-speaking peers. This overall difference was attributable primarily to the relatively lower scores of Spanish speakers on 5 of the 14 measures in the battery, including three tasks (i.e., Identities and Oddities, BVRT matching, and BVRT recognition memory) which generally are considered to be nonverbal. On the language-based tasks (i.e., category fluency and abbreviated BDAE auditory comprehension), issues of word frequency or salience may have contributed to the observed results. For example, the number of exemplars in each semantic category of the category fluency test may not be comparable across languages. The BVRT and Identities and Oddities, however, ostensibly are language-free, suggesting that additional non-language factors also may have contributed to the differences in test performance between English and Spanish speakers.

We had not anticipated the observed pattern of results, particularly the relatively poor performance of Spanish-speaking elders on BVRT matching and recognition memory and Mattis Identities and Oddities. We hypothesize that the geometric nature of the stimuli on these measures may have conferred an advantage to the English speakers that was not shared by the

Spanish speakers, perhaps by virtue of differences in educational emphasis or exposure. (We currently are investigating this theory by evaluating whether Spanish- and English-speaking elders respond differently to geometric and non-geometric stimuli). An alternative factor that may have contributed to the group differences on the BVRT is lack of familiarity by Spanish speakers with the multiple-choice format of this measure; however, if this were the case, we would predict that Spanish speakers would also have scored relatively lower on delayed multiple-choice recognition memory on the SRT, which they did not. Similarly, a differential appreciation of the concept of sameness could result in group performance differences on the BVRT and Identities and Oddities, but such differences would also have been expected on WAIS-R Similarities. A potential explanation for this discrepancy between Similarities and the BVRT and Identities and Oddities, however, is that English and Spanish speakers may apply the concept of sameness differentially to verbal and nonverbal stimuli.

The extent to which our nonverbal measures truly are language-free is an empirical question. Many of the stimuli on the BVRT and Identities and Oddities are geometric and, therefore, have verbal labels. As a result, it is possible that subjects used verbal strategies to perform these tasks, and that this contributed to the observed results. One would expect, however, that if use of verbal mediation on the BVRT and Identities and Oddities contributed to the observed significant group differences on these measures, then group differences should also have been observed on such highly verbal measures as the Boston Naming Test or Selective Reminding Test. Clearly, further investigations are needed to illuminate the factors that influence responses on measures with figural stimuli, such as the BVRT and Identities and Oddities.

Our regression analyses did not serve to clarify the specific factors within the Spanish-speaking group that influenced their performance on cognitive testing. Neither proficiency in English nor duration of time living in the United States was significantly associated with test performance. Having emigrated from Cuba

was associated with better performance on Benton Matching; otherwise, country of origin was not a significant predictor of test performance. Although age and education were associated with performance on several measures within the Spanish-speaking group, age and education could not have contributed to the difference in performance between English and Spanish speakers in the omnibus MANOVA, because groups were matched on these factors.

A measure of acculturation (e.g., Cuellar, Harris, & Jasso, 1980; Marin, Sabogal, VanOss Marin, Otero-Sabogal, & Perez-Stable, 1987) may have further illuminated those characteristics of our Spanish-speaking group that impacted their performance on test measures. Level of acculturation has been shown to affect performance on neuropsychological measures in Hispanic groups (Arnold, Montgomery, Castaneda, & Longoria, 1994). Because no formal measure of acculturation was available in this cohort, we examined the effect of English fluency as an estimate of level of acculturation. Results of age- and education-matched group comparisons suggest that level of acculturation, as reflected by English fluency, was, in fact, an important determinant of test performance. Whereas subjects tested in Spanish who spoke English "not well" or "not at all" scored significantly lower than subjects tested in English on several measures, including two nonverbal tasks, subjects tested in Spanish who reported that they could speak English "well" or "very well" did not differ significantly from subjects tested in English on any measure.

The current study has several advantages relative to previous neuropsychological comparisons of Spanish and English speakers. First, our cohort was randomly selected. Thus, results can be generalized to the population of elders from which the cohort was drawn (i.e., nondemented elders residing in northern Manhattan, New York). One limitation, however, is that our Spanish-speaking cohort is comprised almost entirely of Caribbean Hispanic immigrants. The extent to which our results can be generalized to other Hispanic cultures, such as persons from Mexico, Spain, Central or South America, or persons in their native countries remains to be

determined. A second advantage of the current study is that our sample sizes were larger than many previous studies, resulting in sufficient statistical power to reveal even subtle group differences. Third, our samples of English and Spanish speakers were matched for age and years of education, thereby eliminating these variables as potential confounders in the interpretation of the obtained results. Nevertheless, matching subject groups on years of education does not necessarily equate groups for educational curriculum; hence, differences in educational emphasis may have contributed to the observed results.

As the population of elderly Hispanics continues to increase, so too the need for culturally appropriate instruments to measure cognition in this population increases. Ardila (Ardila, 1995; Ardila, Rosselli, & Puente, 1994) and Loewenstein (Loewenstein, Arguelles, Arguelles, & Linn-Fuentes, 1994), among others, have cautioned that neuropsychologists must be aware of the salience and relevancy of tests and test stimuli to particular cultural groups which they assess. We concur with these authors, and emphasize that caution is warranted when using nonverbal as well as verbal measures in ethnically diverse cohorts. Results of the current study support previous observations (Loewenstein et al., 1993; Lopez & Taussig, 1991) that clearly indicate that using nonverbal tests, in and of itself, does not necessarily eliminate cultural bias in the assessment of Hispanic elders.

Directions for future research include determining *how* Spanish speakers perform differently on neuropsychological tests, and, more important, determining *why* they obtain lower scores on certain neuropsychological measures. Numerous factors influence performance on standardized tests. We were able to examine only a few of these factors (e.g., country of origin, English proficiency, duration of time living in the US) in the current study. Additional variables that impact test performance include environmental contexts, such as learning opportunities and ecological demands; cultural dictates regarding the relevance of different cognitive abilities, test stimuli, and task demands; and nonspecific factors of the testing situation in general (Ardila, 1995). A challenge for cross-

cultural neuropsychology will be to develop objective measures of these factors so that their influences on neuropsychological test scores can be examined empirically.

REFERENCES

- Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, 17, 143-150.
- Ardila, A., Rosselli, M., & Puente, A.E. (1994). *Neuropsychological evaluation of the Spanish speaker*. New York: Plenum.
- Arnold, B.R., Montgomery, G.T., Castaneda, I., & Longoria, R. (1994). Acculturation and performance of Hispanics on selected Halstead-Reitan neuropsychological tests. *Assessment*, 1, 239-248.
- Benton, A.L. (1955). *The Visual Retention Test*. New York, NY: The Psychological Corporation.
- Benton, A.L., & Hamsher, K.deS. (1976). *Multilingual Aphasia Examination*. Iowa City, IA: University of Iowa.
- Bird, H.R., Canino, G., Stipek, M.R., & Shrout, P. (1987). Use of the Mini-Mental State Examination in a probability sample of a Hispanic population. *The Journal of Nervous and Mental Disease*, 175, 731-737.
- Buschke, H., & Fuld, P.A. (1974). Evaluating storage, retention, and retrieval in disordered memory and learning. *Neurology*, 24, 1019-1025.
- Campbell, P.R. (1994). *Current population reports: Population projections for states, by age, sex, race, and Hispanic origin: 1993 to 2020*. Washington, DC: U.S. Department of Commerce.
- Cuellar, I., Harris, L.C., & Jasso, R. (1980). An acculturation scale for Mexican American normal and clinical populations. *Hispanic Journal of Behavioral Sciences*, 2, 199-217.
- Davis, C., Haub, C., & Willett, J. (1983). U.S. Hispanics: Changing the face of America. *Population Bulletin*, 38, 3-43.
- Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). 'Mini-mental State': a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12, 189-198.
- Goodglass, H., & Kaplan, E. (1983). *The assessment of aphasia and related disorders*. Philadelphia, PA: Lea & Febiger.
- Heaton, R.K., Grant, I., & Matthews, C.G. (1986). Differences in neuropsychological test performance associated with age, education, and sex. In I. Grant & K.M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric disorders* (pp. 100-120). New York: Oxford University Press.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia, PA: Lea & Febiger.
- Loewenstein, D.A., Arguelles, T., Arguelles, S., & Linn-Fuentes, P. (1994). Potential cultural bias in the neuropsychological assessment of the older adult. *Journal of Clinical and Experimental Neuropsychology*, 16, 623-629.
- Loewenstein, D.A., Arguelles, T., Barker, W.W., & Duara, R. (1993). A comparative analysis of neuropsychological test performance of Spanish-speaking and English-speaking patients with Alzheimer's disease. *Journal of Gerontology: Psychological Sciences*, 48, 142-149.
- Lopez, S.R., & Taussig, I.M. (1991). Cognitive-intellectual functioning of Spanish-speaking impaired and nonimpaired elderly: Implications for culturally sensitive assessment. *Psychological Assessment*, 3, 448-454.
- Marin, G., Sabogal, F., VanOss Marin, B., Otero-Sabogal, R., & Perez-Stable, E.J. (1987). Development of a short acculturation scale for Hispanics. *Hispanic Journal of Behavioral Sciences*, 9, 183-205.
- Marin, G., & VanOss Marin, B. (1991). *Research with Hispanic populations*. Newbury Park, CA: Sage Publications.
- Mattis, S. (1976). Mental status examination for organic mental syndrome in the elderly patient. In L. Bellak & T.B. Karasu (Eds.), *Geriatric psychiatry* (pp. 77-121). New York, NY: Grune & Stratton.
- Padilla, A.M. (1980). (Ed.) *Acculturation, theory, models, and some new findings*. Boulder, CO: Westview Press for the American Association for the Advancement of Science.
- Rosen, W. (1981). *The Rosen Drawing Test*. Bronx, NY: Veterans Administration Medical Center.
- SPSS for Windows, Release 6.0. (1993) Chicago: SPSS Inc.
- Valle, R., Hough, R., Kolody, B., Cook-Gait, H., Figueroa, G., Jimenez, R., & Atkinson, J.H. (1991). The validation of the Blessed Mental Status Test and the Mini-Mental Status Examination with an Hispanic population. In *Final report to the National Institute of Mental Health, Grant #R01 MH43390-01*.
- Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised*. New York, NY: The Psychological Corporation.
- Zarit, S.H., Miller, N.E., & Kahn, R.L. (1978). Brain function, intellectual impairment and education in the aged. *Journal of the American Geriatric Society*, 26, 58-67.