

UC San Diego

UC San Diego Previously Published Works

Title

Neuropsychological Norms for the U.S.-Mexico Border Region in Spanish (NP-NUMBRS)

Project: Methodology and sample characteristics

Permalink

<https://escholarship.org/uc/item/4vz0g23j>

Journal

The Clinical Neuropsychologist, 35(2)

ISSN

1385-4046

Authors

Cherner, Mariana

Marquine, María J

Umlauf, Anya

et al.

Publication Date

2021-02-17

DOI

10.1080/13854046.2019.1709661

Peer reviewed



Published in final edited form as:

Clin Neuropsychol. 2021 February ; 35(2): 253–268. doi:10.1080/13854046.2019.1709661.

Neuropsychological Norms for the U.S.-Mexico Border Region in Spanish (NP-NUMBRS) Project: Methodology and Sample Characteristics

Mariana Cherner^{a,*}, María J. Marquine^{a,*}, Anya Umlauf^a, Alejandra Morlett Paredes^a, Monica Rivera Mindt^{b,c,d}, Paola Suárez^e, David Yassai-Gonzalez^a, Lily Kamalyan^a, Travis Scott^b, Anne Heaton^a, Mirella Diaz-Santos^e, Amanda Gooding^a, Lidia Artiola i Fortuny^f, Robert K. Heaton^a

^aDepartment of Psychiatry, University of California San Diego

^bDepartment of Psychology, Fordham University

^cDepartment of Latin American and Latina/o Studies Institute, Fordham University

^dDepartment of Neurology, Icahn School of Medicine at Mount Sinai, New York, USA

^eUniversity of California Los Angeles

^fPrivate Practice, Tucson, AZ

Abstract

Objective: The present paper describes the methodology and sample characteristics of the Neuropsychological Norms for the U.S.-Mexico Border Region in Spanish (NP-NUMBRS) Project, which aimed to generate demographically-adjusted norms for a battery of neuropsychological tests in this population.

Methods: The sample consisted of 254 healthy Spanish-speakers, ages 19-60 years, recruited from the U.S.-Mexico border regions of Arizona ($n=102$) and California ($n=152$). Participants completed a comprehensive neuropsychological test battery assessing multiple domains (verbal fluency, speed of information processing, attention/working memory, executive function, learning and memory, visual-spatial skills and fine motor skills). Fluency in both Spanish and English was assessed with performance-based measures. Other culturally-relevant data on educational, social, and language background were obtained via self-report. Demographic influences on test performances were modeled using fractional polynomial equations that allow consideration of linear and non-linear effects.

Results: Participants tested in Arizona had significantly fewer years of education than those tested in California, with no other significant demographic differences by site or cohort. Age and

Corresponding Author: Mariana Cherner, PhD, 9500 Gilman Drive, La Jolla, CA 92093-8231, Phone: 619-823-6340, mcherner@ucsd.edu.

*Joint first authors

Disclosures

In accordance with Taylor & Francis policy and ethical obligation as a researcher, Dr. Heaton is reporting that he receives royalties from the publisher of the Wisconsin Card Sorting Test-64 item, which is one of the tests included in this norming study.

gender were similar across education ranges. Two thirds of the sample were Spanish dominant and the remainder were considered bilingual. Individual articles in this Special Issue detail the generation of demographically adjusted T-scores for the various tests in the battery as well as an exploration of bilingualism effects.

Discussion: Norms developed through the NP-NUMBRS project stand to improve the diagnostic accuracy of neuropsychological assessment in Spanish-speaking young-to-middle-aged adults living in the U.S.-Mexico border region. Application of the present norms to other groups should be done with caution.

Keywords

Cognition; Hispanics; Latinos

Introduction

Normative standards are essential for interpreting neuropsychological test results. It is well understood that person-variables such as age, sex, and exposure to formal education can contribute to normal variation in test performance (Heaton, Ryan, & Grant, 2009). Sociocultural and linguistic factors have more recently also received attention as possible influences (Arnold, Montgomery, Castañeda, & Longoria, 1994; Bure-Reyes et al., 2013). While the complexity and interaction of predictors remains difficult to quantify, it has become increasingly clear that regional or population-specific norms result in better diagnostic accuracy. Faced with the need to evaluate native Spanish speaking adults in our region (Census Bureau, 2017; Ryan, 2013), we undertook a neuropsychological test norming effort targeting the United States (U.S.)-Mexico borderlands of San Diego, CA and Tucson, AZ. In this article, we describe the strategy for participant recruitment and the composition of the resulting sample, the approach to the adaptation of test instructions and stimuli, and the statistical methods used to generate demographically adjusted test norms. Details applicable to the specific tests in our battery appear in the corresponding articles in this issue.

Method

Participants

Participants were 254 adults living in the U.S.-Mexico border region of California and Arizona, who were participants in two cohort studies. The first cohort (Cohort 1, $N=183$) was recruited to develop neuropsychological test norms for Spanish-speakers living in the U.S.-Mexico border region of Tucson, Arizona ($n=102$) and San Diego, California ($n=81$). The second cohort (Cohort 2, $N=71$) comprised healthy control participants from a study of cognitive and functional outcomes among Spanish speakers living with HIV in the U.S.-Mexico border region of San Diego, CA only. Data for Cohort 1 were gathered between 1998 and 2000 and data for Cohort 2 were collected between 2006 and 2009. We combined these cohorts to increase the size of the normative sample. Both studies had similar inclusion/exclusion criteria for healthy controls. In order to participate in either study, volunteers had to be generally healthy women and men between the ages of 18 and 60 who

were native Spanish speakers and who resided in the U.S. at least part of the time. In order to assemble a sample that was more representative of the U.S. borderland population, we did not seek to enroll or test participants in Mexico. Rather, we targeted individuals who lived and/or spent time in the U.S. on a regular basis.

Potential participants were screened for exclusionary conditions using a structured interview to capture any central nervous system disorder or other medical condition (e.g., chronic obstructive pulmonary disease, or significant peripheral injuries or disabilities) that may influence test performance; serious psychiatric conditions including psychosis, bipolar disorder, and severe symptoms of depression or anxiety; and any lifetime substance dependence. Less serious conditions such as hypertension, metabolic disorders, or certain peripheral injuries were reviewed for inclusion by the senior investigators on a case-by-case basis. Participants were also excluded if it was clear at screening that English was their better language. Further language assessments were conducted once enrolled to verify that participants were not English-dominant (see Procedures). Participants were recruited through flyers in community settings frequented by Spanish speakers, as well as in response to in-person presentations by study staff in Hispanic serving community and health-care organizations in San Diego, CA and Tucson, AZ.

Procedures

Test Adaptation.—Table 1 lists the references for the English language version of each test, and for the Spanish language version when available. Test materials not already available in Spanish were adapted. Our approach to the adaptation of test materials was to remain as faithful as possible to the English language version of each test, while being careful to arrive at translations that would be linguistically neutral across the Spanish-speaking world and understood by persons across levels of formal education. With the exception of verbal learning and memory, for which we adapted the test stimuli, and the Paced Auditory Serial Addition Task (PASAT; (Diehr et al., 2003; Gronwall & Sampson, 1974)) for which we recorded the numbers in Spanish, the adaptation was limited to test instructions.

We assembled a group of native Spanish speakers from among our team of psychologists, psychometrists, and PhD students who represented Argentina, Colombia, Cuba, Mexico, Puerto Rico, and Spain. Different team members took the lead in translating each set of instructions, which were then circulated to the rest of the team for feedback. Any words that were idiosyncratic to a specific region, or judged to be of low frequency in the language (e.g., only likely to be recognized by those with high education), were replaced until all were satisfied with the content. That version was then subjected to back translation by an independent reader and checked against the original English version.

Neuropsychological Test Battery.—Participants completed a comprehensive neuropsychological test battery in Spanish assessing domains of verbal fluency, speed of information processing, attention/working memory, executive function, learning and memory, visuospatial, and fine motor skills (see Table 1 for a list of tests by domain). While the majority of the tests were administered to both cohorts, a subset of tests was

administered only to participants in Cohort 1 (i.e., Wechsler Adult Intelligence Scale-Revised [WAIS-R] Arithmetic and Block Design subtests (Wechsler, 1981), and Finger Tapping Test (Halstead, 1947; Reitan & Wolfson, 1993; Spreen & Strauss, 1998). Additionally, some tests were added after the start of data collection for Cohort 1 (WAIS-III Digit Symbol, Symbol Search and Letter-Number Sequencing subtests (Wechsler, 1997), Hopkins Verbal Learning Test-Revised (Brandt & Benedict, 2001), Brief Visuospatial Memory Test-Revised (Benedict, 1997), Wisconsin Card Sorting Test-64 item (Greve, 2001; Kongs, Thompson, Iverson, & Heaton, 2000), resulting in 52 participants from Cohort 1 without data for these tests. There were a few cases for whom data on specific tests were excluded or not collected because of threats to validity, such as poor command of the alphabet for tests that required letter sequencing, examiner error, or unexpected distractions, among others. Table 1 lists the resulting overall *ns* for each test in the battery, and Appendix A lists the *ns* for each test by cohort (Cohort 1 and 2) and by study site (Arizona and California). Importantly, with the majority of participants, the tests were co-normed, facilitating direct comparisons of demographically-corrected standard scores (T-scores) across instruments.

Bilingual study psychometrists in Tucson and San Diego were trained to administer the test battery in a standardized manner. All psychometrists were certified by a senior bilingual psychometrist at the University of California San Diego (UCSD) who also performed quality control on the first 10 protocols administered by each, as well as every 10th protocol thereafter to guard against procedural drift. Each protocol was double-scored by a second psychometrist. If errors were found, they had to be rectified by the original psychometrist prior to data entry. Participants were typically tested in a research setting (Dr. Heaton's lab at the HIV Neurobehavioral Research Center, UCSD) or a clinical setting (Dr. Artiola's office in Tucson). On rare occasion, participants for whom that was inconvenient were tested in their home.

Educational, Social, and Linguistic Background.—Participants completed self-report questionnaires aimed at collecting information on culturally-relevant background information. Items regarding educational background covered total years of formal schooling, including years of education in the U.S. and in the country of origin, reasons for discontinuing attending school, and details on the elementary school attended in the participants' country of origin. Specifically, participants were asked to indicate whether the school was large (a large school that had many classrooms and room to play), regular (a school of regular size that had at least one classroom per grade and room to play), or small (a small school with less than one classroom per grade); and to report the typical number of students in a class (less than 10, 11 to 20, 21 to 30, 31 to 40, or more than 40). Regarding social background, participants reported the number of years of education completed by their parents, years participants spent living in the U.S. and their country of origin, childhood socioeconomic status, whether they worked as a child, and whether they were currently gainfully employed. We assessed language use in order to exclude any participants whose better language was English. Language preference across a number of daily life activities (e.g., listening to the radio, watching TV, reading, speaking with family and friends, praying) was measured utilizing a scale from 1 "Always in Spanish" to 5 "Always in

English”, with 3 being “similarly in English and Spanish”. As detailed in the article by Suárez et al. (this issue), participants also completed the Controlled Oral Word Association Test with letters F-A-S in English (Strauss, Sherman, & Spreen, 2006) and P-M-R in Spanish (Artiola i Fortuny, Hermosillo Romo, Heaton, & Pardee, 1999; Strauss et al., 2006). We estimated English fluency as the ratio of FAS to total words in both languages [FAS/(FAS+PMR)] (Suárez et al., 2014) and classified as English-dominant those with scores .67 (i.e., greater than or equal to 2/3 of all words in English). Any participants in this range were excluded from analyses. Participants with scores less than or equal to .33 were considered monolingual Spanish-speakers, and those in between were considered bilingual. These scores were computed for both Cohort 1 and Cohort 2. However, for Cohort 2, the majority of FAS scores were mistakenly discarded after participants were classified as monolingual or bilingual at their screening visit for the purposes of that study. As a result, while we are able to describe the proportion of participants judged to be bilingual, English fluency scores for these particular participants are unfortunately not available to report or include in analyses. We recognize that this simplified quantification of bilingualism does not capture all the nuances of this construct, but at the time of data collection, our primary purpose had been to avoid enrolling primary English speakers, rather than in-depth characterization of bilingualism or its effects.

In order to standardize the coding of educational attainment to be used for demographic adjustments, we employed the guidelines advocated by Dr. Robert Heaton, which were employed for the Expanded Halstead-Reitan Battery (Heaton, Miller, Taylor, & Grant, 2004) and the Batería Neuropsicologica en Español (Artiola i Fortuny et al., 1999) standardization samples. The possible range of education is 0 (less than one year or no formal schooling) to 20 (doctoral degree). There are two general principles: (1) Give credit for the highest grade or degree completed, not for the number of years it may have taken to obtain, and not for partial years. For example, if someone repeated the second grade and subsequently graduated from high-school, they would be assigned 12 years, even though it took 13 to complete. If someone dropped out of school part-way through the 6th grade, they would be assigned 5 years. If someone completed two separate PhDs, they would still receive 20 years. (2) Only academic school-based years count. Vocational training (e.g., auto mechanics certification) or post-degree training (e.g., medical residency) do not add to the years of education. Because education influences performance on the majority of tests, it is critically important that years of formal schooling be entered into our normative equations in a standard manner. For this reason, clinicians or researchers applying our norms need to be knowledgeable about the number of years that are associated with various educational milestones, and also need to be able to differentiate between academic and vocational programs, in this case, for persons educated in Mexico. We used the rubric shown in Table 2 to standardize the assignment of years of formal schooling, based on the Mexican and U.S. educational systems.

Statistical Approach

After performing routine value range checks to guard against data entry errors, there was a need to convert the raw scores to a uniform metric. This is because tests of cognitive abilities are measured on different scales, with different distributions. Some may measure the time it

takes to complete a task (e.g., the Trail Making Test, (Army Individual Test Battery, 1944) or the number of errors made (e.g., Halstead Category test, (DeFilippis & McCampbell, 1979; Reitan & Wolfson, 1993)), where low scores correspond to better performance on a test, whereas others may measure the total number of correct responses (e.g., WAIS-III Letter-Number Sequencing subtest (Wechsler, 1997)), with high scores corresponding to better performance. To impose uniformity on the test results, raw scores were converted into normally distributed scaled scores by measuring standardized quantiles of the raw scores and scaling them to have the mean of 10 and standard deviation (*SD*) of 3, such that higher values always correspond to better performance on the test. In the second step of the normative procedure, the scaled scores were regressed on age, education, and sex – the demographic characteristics that frequently are correlated with cognitive abilities – using a multiple fractional polynomial (MFP) procedure (Royston & Altman, 1994). This method allows for inclusion of both linear and non-linear effects for numeric predictors (i.e., age and education) and selects the best curve ($p < 0.05$) from several options: linear, quadratic, logarithmic, and other combinations of fractional polynomials of first (e.g. x^m) and second degree (e.g. $x^{m1} + x^{m2}$) with powers (m_i) ranging from -2 to $+3$. Residuals, calculated as differences between the recorded scores and scores predicted by the MFP model, were standardized and scaled (mean=50, *SD*=10). These calculated values are termed demographically corrected T-scores.

Sensitivity analyses were performed for each test at multiple stages. First, the stability of the fractional polynomial, used in calculations of T-scores, was assessed as recommended (Royston & Sauerbrei, 2009). A bootstrap ($K=1000$) method was used to sample with replacement from the data and an MFP model was fitted each time. The frequency of the original polynomial in the bootstrap samples was assessed separately for each numeric predictor (age and education). In addition, the bootstrap procedure was used to generate ‘bagged’ estimate of the MFP curve (and its 95% confidence boundaries), which was then compared to the curve fitted on the original data. Both assessments were done to see if the curve obtained by the norming procedure is really the best fitting curve or a curve obtained by chance or abnormalities in the data (e.g., outliers). Secondly, we tested to see if the application of normative formulas resulted in T-scores free of demographic effects. Associations between T-scores and demographic characteristics were tested with the t-test for two independent samples (sex effect) and with Pearson’s correlation test (age and, separately, education effects). The norming procedure was considered successful in removing demographic effects if the p-values for the above associations were greater than 0.2. In the third step of sensitivity analysis, normative formulas were applied to hypothetical data to test the results of extrapolation; application of formulas to a person with a combination of test scores and demographic characteristics not seen in the normative data. All extreme results, i.e. T-scores that are extremely low or extremely high (± 5 *SD*), were investigated to see if they were due to an unusual and unlikely combination of parameters (e.g., poor test performance in a younger person with college education) or due to, possibly, a poor model fit. All computations and statistical test procedures were performed using R software (R Core Team, 2018) and “mfp” R package (Benner, 2005).

The articles in this issue describing the development of norms for each test all contain the following essential information: (1) descriptive characteristics and distribution of raw scores

for each test; (2) univariable and two-way-interactive association of demographics characteristics (age, education, and gender) with test raw scores; (3) for test measures that met sensitivity criteria described above, we report Scaled Scores and demographically-adjusted T-scores; and (4) for tests with limited range of scores or very skewed raw score distribution, we report percentile ranges. In addition, (a) we provide information on the distribution of the resulting T-scores (M, SD, skewness, range); (b) we confirmed that the resulting T-scores were free from demographic effects, by investigating their univariable association with age, education, and gender; (c) we tested for cohort effects by comparing the resulting T-scores between participants in Cohort 1 and Cohort 2, as well as between participants tested in Arizona and California via independent samples t-tests; and (d) we calculated the T-scores that would result by applying published norms for English-speaking non-Hispanic Whites and non-Hispanic Blacks/African Americans in the U.S. when available. Using McNemar's test allows us to compare the expected rates of neurocognitive performances that are below expectations (defined as T-scores <40, i.e., below 1 standard deviation), based on the normal distribution, to those that would be obtained in our normative sample if we apply existing norms for English-speaking non-Hispanics.

Results

Characteristics of the Normative Sample

Table 3 shows the demographic characteristics (age, education, and gender) of the overall normative sample, as well as by study cohort (Cohort 1 and 2) and by study site (Arizona or California), and Figures 1a and 1b show the distributions of age and education in the overall sample, respectively. Participants were generally evenly distributed across years of age from 19 to 60, except that there were relatively fewer participants in the older age ranges. Education spanned from 0 to 20 years, with relatively fewer participants at the ends of the distribution. A series of independent sample t-tests and chi-square tests showed no significant differences in age ($p=.69$) or gender ($p=.70$) between Cohorts 1 and 2, but Cohort 1 had significantly fewer years of formal education than Cohort 2 ($p<.0001$; Table 3). This is primarily because Cohort 1 was recruited specifically to represent the full range of educational attainment, whereas Cohort 2 was enrolled as controls for a case-control study in San Diego. Analyses in the subset of participants tested in San Diego only ($n=152$), yielded comparable findings, i.e., Cohort 1 had lower formal education than participants from Cohort 2, and there were no significant differences in age ($p=.87$) or gender ($p=.35$) based on the two time spans when the data were collected (see Appendix B, $p<.0001$). There were no significant differences between participants tested in Arizona and California on any demographic factors in the overall cohort (Table 3; $ps>.26$) or among participants from Cohort 1 only (Appendix B, $ps>.16$). Within participants from Cohort 1 (Appendix B), there were also no significant differences on age ($p=.12$), education ($p=.40$), or gender ($p=.97$) between those who completed all tests ($n=131$) and the subset who completed only a subset of the tests ($n=52$).

Table 4 shows demographic characteristics of the overall normative sample ($N=254$) by level of education, stratified for descriptive purposes only (6 years, 7 to 10 years, 11 to 12 years

13 years). There were no significant differences in age ($F[3, 250] = 2.12, p=.10$), or gender distribution ($Chi-Square [3] = 2.96, p=.39$) among education ranges.

Table 5 lists educational, social, and language background characteristics of the study sample. A majority of participants completed more years of education in their country of origin than in the U.S., with 63.72% of the sample completing all of their education in their country of origin. Most participants attended a large school with class sizes of 21 to 30 students, and almost a third of the sample had to stop attending school to work. Years of education completed by both parents was on average 6-7. Participants lived a majority of their lives in their country of origin, with an average of nearly 11 years in the U.S. Most participants described their childhood socioeconomic status as middle class, but nearly a third reported having been poor. Approximately half of participants reported working for money during childhood, with 22.84% of the sample working at age 10 or younger. Over one third of those who reported working as a child stated they did so to help their families financially. Over two thirds of participants were gainfully employed at the time of their participation in the present study. All but one of the participants reported that Spanish was the first language they learned. Two-thirds of the sample was monolingual Spanish-speaking or strongly Spanish dominant, with the remaining third being bilingual. Average ratings of language used in various everyday activities (listening to the radio or watching TV, reading, doing math calculations, praying and conversing with family) indicated that Spanish was the predominant language used in daily life.

Discussion

In this special issue, we provide demographically adjusted norms for Spanish speakers from the U.S.-Mexico border regions of California and Arizona for a battery of tests that have the advantage of co-norming within the same population, as well as having undergone identical methods of data analysis. There are important limitations that need to be considered.

Our sample size is modest for norms generation purposes. Larger samples can be expected to increase the reliability of T-scores and estimates of demographic effects. Nevertheless, our sample size is considered sufficient to accommodate modeling the demographic effects (Harris, 1985; Maindonald & Braun, 2007).

The resulting norms are applicable to adults age 19-60. We do not recommend applying our formulas to ages beyond this range, as the resulting T-scores cannot be considered reliable. Until they are investigated in other Spanish-speaking groups, the norms are considered specific to the U.S.-Mexico border regions of California and Arizona. The article by Morlett-Paredes and colleagues in this issue (Morlett-Paredes et al., under review) lists other norms that may be better suited to different populations, including older adults.

We did not consider effects of bilingualism in the normative equations. U.S.-dwelling Spanish speakers will have varying degrees of bilingualism, and, as described in the article by Suárez and colleagues (P. Suárez et al., under review), English proficiency may affect test performance in the native language, or it may be a proxy for other sociodemographic influences on test performance. In this study, we assessed language preference by self-report

and quantified English language fluency using the COWAT in order to exclude participants who were English-dominant. An unfortunate mishap with our data recording precluded us from investigating the value of incorporating effects of English fluency into the norms. We opted not to model bilingualism as a categorical variable based on the COWAT results because only one third of the sample met that definition, and within both “monolinguals” and “bilinguals” there is still a range of second language proficiency. Similarly, we opted not to model self-reported language use preference as an indicator of bilingualism because of the impracticality of requiring others to use the same instrument we used in order to apply the norms and because despite a range of English proficiency, all participants expressed a preference to be tested in Spanish. Nevertheless, our norms represent values for a group of native Spanish speakers with a range of English fluency from monolingual Spanish-speaking to balanced bilingual. The paper by Suárez et al. in this issue describes effects of bilingualism on our battery in a subset of participants with English fluency data available.

Similarly, we have not accounted for possible effects of other sociodemographic predictors. As shown in Table 4, we collected substantial detail about the life circumstances of our participants, some of which may affect test performance, and some of which may already be captured by accounting for years of formal schooling. As of the publication of this issue, we are still exploring which of the life influences that we captured may be unique and which may share variance in their predictive value. Our sample size may not be sufficient to explore further granularity for all variables.

Despite these limitations, we feel that we are making an important contribution by providing a co-normed test battery that was generated through careful participant characterization and rigorous statistical analysis that included modeling years of education as a continuous variable and exploring non-linear effects of demographic predictors. The articles that follow detail the methods for the various tests and ability domains in the NP-NUMBRS battery.

Acknowledgements

This work was supported by grants from the National Institutes of Health (P30MH62512, R01MH064907, R01MH57266, K23MH105297, P30AG059299) and the UCSD Hispanic Center of Excellence (funded by the Health Resources & Services Administration grant D34HP31027).

Appendix

Appendix A.

Number of participants with data on each test of the neuropsychological test battery by study cohort and site

Domain	Test	Total <i>N</i>	Study Cohort		Study Site	
			Cohort 1	Cohort 2	AZ	CA
			<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
Verbal Fluency	Letter Fluency	254	183	71	102	152
	Animal Fluency	254	183	71	102	152

Domain	Test	Total <i>N</i>	Study Cohort		Study Site	
			Cohort 1	Cohort 2	AZ	CA
			<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
Speed of Information Processing	Trail Making Test A	251	182	69	102	149
	WAIS-III Digit Symbol	201	131	70	62	139
	WAIS-III Symbol Search	200	131	69	62	138
Attention/Working Memory	PASAT-50	249	182	67	101	148
	WAIS-III L-N Sequencing	202	131	71	62	140
	WAIS-R Arithmetic	178	178	0	102	80
Executive Function	Trail Making Test B	246	176	70	99	147
	WCST-64	189	120	69	59	130
	Halstead Category Test	252	183	69	102	150
Learning and Memory	Hopkins Verbal Learning Test-Revised	202	131	71	62	140
	Brief Visuospatial Memory Test – Revised	202	132	70	63	140
Visual-spatial skills	WAIS-R Block Design	181	181	0	101	80
Fine motor skills	Grooved Pegboard	254	183	71	102	152
	Finger Tapping	183	183	0	102	81

Note. AZ = Arizona; CA=California

Appendix B.

Demographic characteristics by cohort subsets

Characteristic	Cohort by Time (San Diego only)		Study Site (Cohort at Time 1 only)		Tests completed (Cohort at Time 1 only)	
	Cohort 1 (<i>n</i> =81)	Cohort 2 (<i>n</i> =71)	Arizona (<i>n</i> =102)	California (<i>n</i> =81)	All tests (<i>n</i> =131)	Subset of tests (<i>n</i> =52)
Age, <i>M</i> (<i>SD</i>) [range]	38.02 (8.46) [21-55]	37.75 (11.82) [19-60]	36.41 (10.39) [20-55]	38.02 (8.46) [21-55]	37.84 (9.48) [20-55]	35.33 (9.72) [20-55]
Education, <i>M</i> (<i>SD</i>) [range]	9.43 (4.44) [0-20]	12.59 (4.13) [2-20]	10.32 (3.98) [2-20]	9.43 (4.44) [0-20]	9.77 (4.31) [0-20]	10.33 (3.92) [2-18]
Gender (Female), <i>N</i> (%)	43 (53.09)	38 (46.91)	63 (61.76)	43 (53.09)	76 (58.02)	30 (57.69)

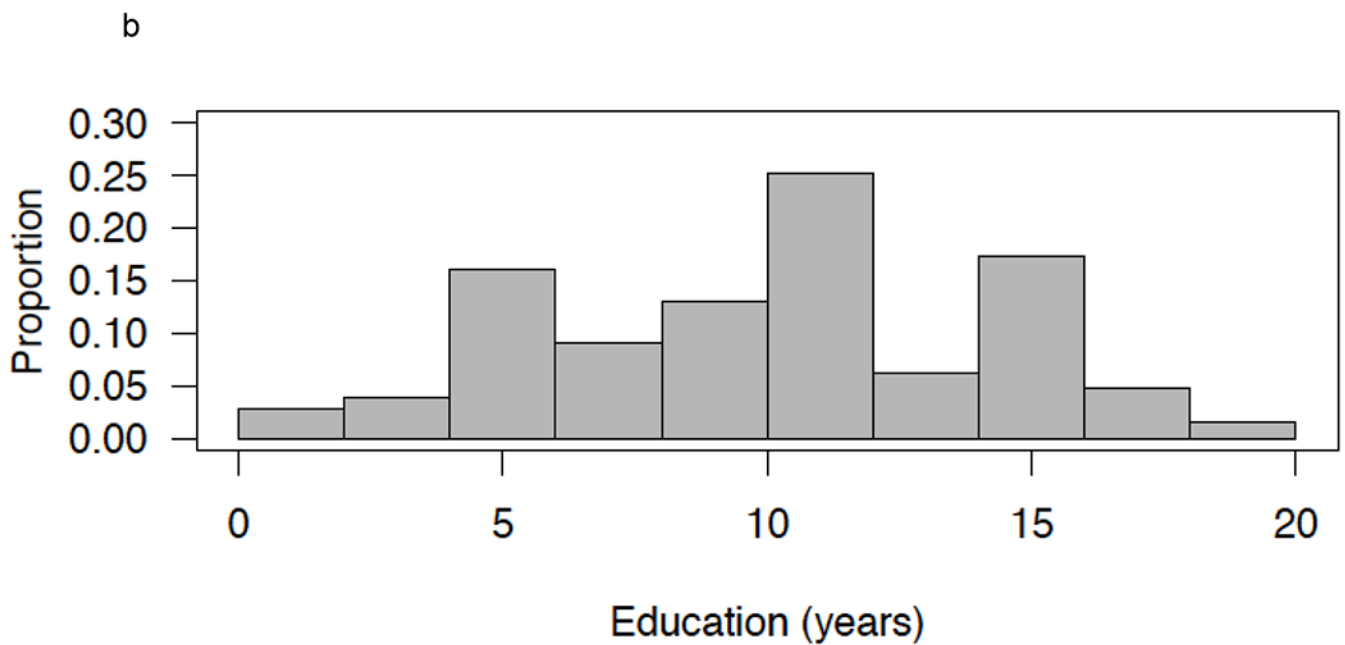
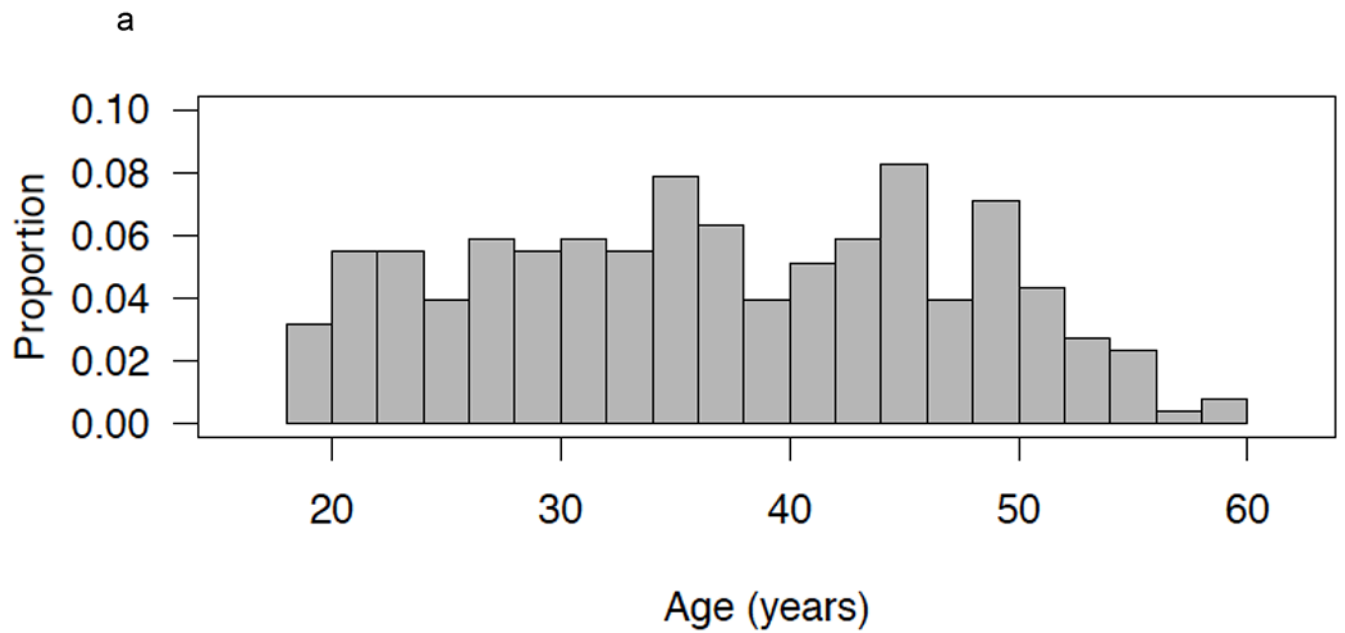
Note. *M* = mean; *SD* = standard deviation

References

- Army Individual Test Battery. (1944). Manual of directions and scoring. Washington D.C.: Adjunct General's Office, War Department.
- Arnold BR, Montgomery GT, Castañeda I, & Longoria R (1994). Acculturation and performance of Hispanics on selected Halstead-Reitan neuropsychological tests. *Assessment*, 1(3), 239–248.
- Artiola i Fortuny L, Hermosillo Romo D, Heaton RK, & Pardee RE (1999). Manual de Normas y Procedimientos para la Bateria Neuropsicologica en Espanol. Tucson, AZ: m Press.

- Benedict R (1997). Brief Visuospatial Memory Test- Revised. Odessa, FL: Psychological Assessment Resources, Inc.
- Benner A (2005). Multivariable Fractional Polynomials. *R News*(5(2)), 20–23.
- Benton AL, Hamsher K, & Sivan AB (1994). Multilingual Aphasia Examination: Manual of instructions. Iowa City, IA: AJA Associates.
- Brandt J, & Benedict R (2001). Hopkins Verbal Learning Test- Revised. Lutz, FL: Psychological Assessment Resources, Inc.
- Bure-Reyes A, Hidalgo-Ruzzante N, Vilar-Lopez R, Gontier J, Sanchez L, Perez-Garcia M, & Puente AE (2013). Neuropsychological test performance of Spanish speakers: is performance different across different Spanish-speaking subgroups? *J Clin Exp Neuropsychol*, 35(4), 404–412. doi:10.1080/13803395.2013.778232 [PubMed: 23496164]
- Census Bureau, U. (2017). Facts for Features: Hispanic Heritage Month 2017. Retrieved from <https://www.census.gov/newsroom/facts-for-features/2017/hispanic-heritage.html>
- Cherner M, Suarez P, Lazzaretto D, Fortuny LA, Mindt MR, Dawes S, ... group, H. (2007). Demographically corrected norms for the Brief Visuospatial Memory Test-revised and Hopkins Verbal Learning Test-revised in monolingual Spanish speakers from the U.S.-Mexico border region. *Arch Clin Neuropsychol*, 22(3), 343–353. doi:10.1016/j.acn.2007.01.009 [PubMed: 17293078]
- DeFilippis NA, & McCampbell E (1979). Manual for the booklet category test. Odessa, FL: Psychological Assessment Resources.
- Diehr MC, Cherner M, Wolfson TJ, Miller SW, Grant I, Heaton RK, & Group, H. I. V. N. R. C. (2003). The 50 and 100-item short forms of the Paced Auditory Serial Addition Task (PASAT): demographically corrected norms and comparisons with the full PASAT in normal and clinical samples. *J Clin Exp Neuropsychol*, 25(4), 571–585. doi:10.1076/jcen.25.4.571.13876 [PubMed: 12911108]
- Greve KW (2001). The WCST-64: A standardized short-form of the Wisconsin Card Sorting Test. *Clinical Neuropsychologist*, 15(2), 228–234. doi:10.1076/Clin.15.2.228.1901
- Gronwall DMA, & Sampson H (1974). The psychological effects of concussion. New Zealand: Auckland University Press/Oxford University Press.
- Halstead WC (1947). Brain and intelligence. Chicago, IL: University of Chicago Press.
- Harris RJ (1985). A primer of multivariate statistics (2nd Ed.). Section 2.3. New York: Academic Press.
- Heaton RK, Miller SW, Taylor MJ, & Grant I (2004). Revised comprehensive norms for an expanded Halstead-Reitan Battery: Demographically adjusted neuropsychological norms for African American and Caucasian adults. Lutz, FL: Psychological Assessment Resources.
- Heaton RK, Ryan L, & Grant I (2009). Demographic influences and use of demographically corrected norms in neuropsychological assessment. *Neuropsychological assessment of neuropsychiatric and neuromedical disorders*, 3, 127–155.
- Klove H (1963). Clinical Neuropsychology. *Med Clin North Am*, 47, 1647–1658. [PubMed: 14078168]
- Kongs SK, Thompson LL, Iverson GL, & Heaton RK (2000). Wisconsin Card Sorting Test-64 Card Version: Professional Manual. Odessa, FL: Psychological Assessment Resources.
- Maindonald J & Braun J. (2007) Data Analysis and Graphics Using R: An Example-based Approach (2nd edition). p. 199 Cambridge University Press.
- Morlett-Paredes A, Carrasco J, Kamalyan L, Cherner M, Umlauf A, Rivera Mindt M, ... Marquine MJ (under review). Demographically adjusted normative data for the Halstead Category Test in a Spanish-speaking adult population: results from the neuropsychological norms for the US-Mexico border region in Spanish. *The Clinical Neuropsychologist*.
- R Core Team. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Retrieved from <https://www.R-project.org/>
- Reitan RM, & Wolfson D (1993). The Halstead-Reitan neuropsychological test battery: Theory and clinical interpretation (2nd edition) (2 ed.). Tucson, AZ: Neuropsychological Press.

- Royston P, & Altman DG (1994). Regression using fractional polynomials for continuous covariates: Parsimonious parametric modeling. *Journal of the Royal Statistical Society, Series C* 43(3), 429–467.
- Royston P, & Sauerbrei W (2009). Bootstrap assessment of the stability of multivariable models. *The Stata Journal*, 9(4), 547–570.
- Ryan C (2013). Language Use in the United States: 2011. American Community Survey Reports Retrieved January 23, 2015 <http://www.census.gov/prod/2013pubs/acs-22.pdf>
- Spree O, & Strauss E (1998). A compendium of neuropsychological tests (2 ed.). New York, NY: Oxford University Press.
- Strauss E, Sherman EM, & Spree O (2006). A compendium of neuropsychological tests: Administration, norms, and commentary. American Chemical Society.
- Suárez P, Díaz-Santos M, Marquie MJ, Gollan T, Artiola i Fortuny L, Heaton RK, & Cherner M (under review). Role of English fluency on verbal and non-verbal neuropsychological tests in native Spanish speakers from the U.S.-Mexico border region using demographically corrected norms. *The Clinical Neuropsychologist*.
- Suárez PA, Gollan TH, Heaton R, Grant I, Cherner M, & Group H (2014). Second-language fluency predicts native language stroop effects: evidence from Spanish-English bilinguals. *J Int Neuropsychol Soc*, 20(3), 342–348. doi:10.1017/S1355617714000058 [PubMed: 24622502]
- Wechsler D (1981). WAIS-R Manual: Wechsler adult intelligence scale- revised. New York, NY: Psychological Corporation.
- Wechsler D (1997). WAIS-III administration and scoring manual. San Antonio, TX: The Psychological Corporation.
- Wechsler D (2003). Escala Wechsler de Inteligencia para adultos- III. Mexico City, Mexico: Manual Moderno.



Figures 1a and 1b.

Distribution of age (Figure 1a) and education (Figure 1b) in the current sample (N=254).

Table 1.

Neuropsychological test battery

Domain	Test	English Reference	Adaptation to Spanish	Total N
Verbal Fluency	Letter Fluency	(Benton, Hamsher, & Sivan, 1994)	(Ariola i Fortuny et al., 1999)	254
	Animal Fluency	(Benton et al., 1994)	NP-NUMBRS	254
Speed of Information Processing	Trail Making Test A	(Army Individual Test Battery, 1944)	NP-NUMBRS	251
	WAIS-III Digit Symbol	(Wechsler, 1997)	(Wechsler, 2003)	201
	WAIS-III Symbol Search	(Wechsler, 1997)	(Wechsler, 2003)	200
Attention/Working Memory	PASAT-50 and 200	(Diehr et al., 2003);(Gronwall & Sampson, 1974)	NP-NUMBRS	249
	WAIS-III L-N Sequencing	(Wechsler, 1997)	(Wechsler, 2003)	202
	WAIS-R Arithmetic	(Wechsler, 1981)	NP-NUMBRS	178
Executive Function	Trail Making Test B	(Army Individual Test Battery, 1944)	NP-NUMBRS	246
	WCST-64	(Greve, 2001); (Kongs et al., 2000)	(Ariola i Fortuny et al., 1999)	189
	Halstead Category Test	(DeFilippis & McCampbell, 1979); (Reitan & Wolfson, 1993)	NP-NUMBRS	252
Learning and Memory	Hopkins Verbal Learning Test-Revised	(Brandt & Benedict, 2001)	(Cherner et al., 2007)	202
	Brief Visuospatial Memory Test – Revised	(Benedict, 1997)	(Cherner et al., 2007)	202
Visual-spatial skills	WAIS-R Block Design	(Wechsler, 1981)	NP-NUMBRS	181
Fine motor skills	Grooved Pegboard	(Klove, 1963)	NP-NUMBRS	254
	Finger Tapping	(Halstead, 1947); (Reitan & Wolfson, 1993); (Spreen & Strauss, 1998)	NP-NUMBRS	183

Note. NP-NUMBRS = Neuropsychological Norms for the U.S.-Mexico Border Region in Spanish. Notes tests that were adapted into Spanish for the current project.

Table 2.

Guidelines for standardized assignment of years of education

	Years/Milestones	Level by Educational System	
	Completed	Mexico	US
0	Less than first grade		
1	Completed 1 year	Primaria	Elementary
2	Completed 2 years	Primaria	Elementary
3	Completed 3 years	Primaria	Elementary
4	Completed 4 years	Primaria	Elementary
5	Completed 5 years	Primaria	Elementary
6	Graduated - Elementary	Primaria	Elementary/Middle school *
7	Completed 1 year	Secundaria	Middle (Junior High) school
8	Completed 2 years / Graduated Middle	Secundaria	Middle (Junior High) school
9	Graduated – Secundaria/ 1 year High School	Secundaria	High school
10	Completed 2 years Preparatoria/High School	Preparatoria/ Bachillerato	High school
11	Completed 3 years Preparatoria/High School	Preparatoria/ Bachillerato	High school
12	Graduated – Preparatoria/High school	Preparatoria/ Bachillerato	High school
13	Completed 1 year of college/university	Universidad	College
14	Associates degree /2 years college/university	Universidad	College; 2-year college degree
15	Completed 3 years of college/university	Universidad	College
16	Bachelor's degree, Licenciatura ^a	Título universitario de 4 años	College; 4-year university degree
17	Completed 1 year graduate education	Cursando título superior	Enrolled in Master; PhD
18	Licenciatura ^a MA, MS, MEd	Maestria Licenciado en Derecho (Law) Medico Dentista Veterinario	Master's degree
19	JD ^b , EdS		Jurisprudence Doctor (Law); Education Specialist
20	PhD, MD ^b , EdD, DDS, DVM	Doctorado	Doctorate

GUIDELINES:

Assign years completed (not partial) until graduation from High School / Preparatoria (maximum = 12)

* in U.S. some Middle schools start at grade 6, some at grade 7 – always assign years completed

1 full year at a two-year college or four-year university = 13

2 full years at two-year college but no degree = 13

Completed 2-year college degree (e.g., Associate degree) = 14

2 full years at a four-year university = 14

3 or more full years at a four-year university, but no degree = 15

Bachelor's (University) Degree (e.g., BA or BS) = 16 – do not confuse with Bachillerato (High School =12)

^aLicenciatura: Typically equivalent to 4-year college in the U.S., but can be up to 6 years post Preparatoria depending on specialty (e.g., Psychology = 4; Law = 6) = 16-18; assign number of years completed toward degree; if coursework completed but degree not obtained assign maximum for that specialty minus 1.

More than four years to obtain University Degree = 16

^bMedical and Law degree in Mexico = 18 (6 years of university).

1 full year toward Master's Degree = 17 (if also has Licenciatura = 18 years, assign 18)

At least 2 full years toward Master's Degree, but no M.S. = 17 (if Licenciatura = 18 years, assign 18)

Completed Master's Degree (M.A. or M.S. Med) = 18 (if Licenciatura = 18 years, only assign 18)

Enrolled in doctoral program: number of years completed post previous degree (16-19; also see Licenciatura above)

All coursework toward doctorate, but no dissertation/degree = 19

Jurisprudence Doctor (JD) U.S. = 19 - Note: Licenciatura en Derecho in Mexico = 18

Completed doctorate: PhD., MD, DDS, DMV = 20 in US; Note: medical, dental, veterinary degrees in Mexico = 18

Maximum years of education = 20

Diplomado or post-degree specialization does not count toward additional years

Table 3.

Demographic characteristics of the overall study cohort, and of the cohort by years when the data were collected (Cohort 1 and Cohort 2) and study site (Arizona and California)

Characteristic	Total (N=254)	Cohort by Time		Study Site	
		Cohort 1 (n=183)	Cohort 2 (n=71)	Arizona (n=102)	California (n=152)
Years of test	1998-2009	1998-2000	2006-2009	1998-2000	1998-2009
Age					
<i>M</i> (<i>SD</i>)	37.32 (10.24)	37.13 (9.59)	37.75 (11.82)	36.41 (10.39)	37.89 (10.13)
Range	19-60	20-55	19-60	20-55	19-60
<i>N</i> (%)					
19-29	65 (25.59)	45 (24.59)	20 (28.17)	29 (28.43)	36 (23.68)
30-39	82 (32.28)	62 (33.88)	20 (28.17)	32 (31.37)	50 (32.89)
40-49	71 (27.95)	58 (31.69)	13 (18.31)	29 (28.43)	42 (27.63)
50-60	36 (14.17)	18 (9.84)	18 (25.35)	12 (11.77)	24 (15.79)
Education					
<i>M</i> (<i>SD</i>)	10.67 (4.34)	9.93 (4.20)	12.59 (4.13)	10.31 (3.99)	10.91 (4.56)
Range	0-20	0-20	2-20	2-20	0-20
<i>N</i> (%)					
6	58 (22.84)	50 (27.32)	8 (11.27)	24 (23.53)	34 (22.37)
7-10	56 (22.05)	46 (25.14)	10 (14.09)	26 (25.49)	30 (19.74)
11-12	64 (25.19)	45 (24.59)	19 (26.76)	26 (25.49)	38 (25.00)
13	76 (29.92)	42 (22.95)	34 (47.89)	26 (25.49)	50 (32.89)
Female, <i>N</i> (%)	149 (58.66)	106 (57.92)	43 (60.56)	63 (61.76)	86 (56.58)
Right H, <i>N</i> (%) ^a	222 (97.37%)	163 (97.02)	59 (98.33%)	85 (96.9%)	137 (97.86%)

Note.

^a n=228. *M* = mean; *SD* = standard deviation.

Table 4.Demographic characteristics of the normative sample by years of education ($N=254$)

	6 (n=58)	7-10 (n=56)	11-12 (n=64)	13 (n=76)
Age (years), $M(SD)$	39.71 (9.86)	36.95 (9.54)	35.14 (10.34)	37.61 (10.69)
Education (years), $M(SD)$	4.72 (1.55)	8.59 (0.91)	11.81 (0.39)	15.79 (1.67)
% Female	62.07%	55.36%	65.53%	52.63%

Note. M : mean; SD : standard deviation.

Table 5.**Educational, Social, and Language Background Characteristics**

	Descriptive Characteristics	
	<i>M(SD), %</i>	<i>n</i>
Educational Background		
Years of education in country of origin	8.53 (4.80)	227
Years of education in the U.S.	2.53 (4.73)	227
Proportion of education by country		227
More years of education in country of origin	84.14%	191
More years of education in the U.S.	14.98%	34
Equal number of years of education in both countries	0.88%	2
Type of school attended ^a		243
Large	55.56%	135
Regular	39.92%	97
Small	4.53%	11
Number of students in the class		247
Less than 21	15.39%	38
21 to 30	39.27%	97
31 to 40	24.29%	60
40+	21.05%	52
Had to stop attending school to work		224
Yes	28.57%	64
Social Background		
Mother's years of education	5.76 (3.65)	180
Father's years of education	6.79 (5.06)	163
Years lived in country of origin	26.41 (12.49)	245
Years living in the U.S.	10.69 (10.85)	245
Childhood SES		251
Very poor	5.98%	15
Poor	27.09%	68
Middle class	58.17%	146
Upper class	8.77%	22
Worked as a child		248
Yes	52.82%	131
Reason to work		131
Help family financially	38.17%	50
Own benefit	61.83%	81
Age started working as a child	12.98 (3.18)	131
Currently Gainfully Employed		224
Yes	68.75%	154
Language		
First Language		250

	Descriptive Characteristics	
	<i>M(SD), %</i>	<i>n</i>
Spanish	98.40%	246
English	0.40%	1
Both	1.20%	3
Current Language Use Rating ^b		251
Radio or TV	2.37 (1.03)	--
Reading	2.24 (1.18)	--
Math	1.54 (1.05)	--
Praying	1.26 (0.72)	--
With family	1.56 (0.89)	--
Performance-based language fluency		203
Spanish dominant	62.07%	126
English dominant	0.00%	0
Bilingual	37.93%	77

Note. *M*: mean; *SD*: standard deviation.

^aType of school attended: 'large' refers to large school that had many classrooms and room to play; 'regular' refers to a school of regular size that had at least one classroom per grade and room to play; and small school refers to a small school with less than one classroom per grade.

^bRatings for each activity ranged from 1 "Always in Spanish" to 5 "Always in English", with 3 being "similarly in English and Spanish").