# Demographically corrected norms for African Americans and Caucasians on the Hopkins Verbal Learning Test–Revised, Brief Visuospatial Memory Test–Revised, Stroop Color and Word Test, and Wisconsin Card Sorting Test 64-Card Version

Marc A. Norman [a] , David J. Moore [a] , Michael Taylor [a c] , Donald Franklin Jr. [a] , Lucette Cysique [a b] , Chris Ake [a] , Deborah Lazarretto [a] , Florin Vaida [a] , Robert K. Heaton [a] & the HNRC Group

[a] Department of Psychiatry , University of California , San Diego, CA, USA

[b] Brain Sciences, University of New South Wales , Sydney, NSW, Australia

[c] Department of Psychology , San Diego State University , San Diego, CA, USA
Published online: 04 May 2011.

PLEASE SCROLL DOWN FOR ARTICLE

www.manaraa.com

# Demographically corrected norms for African Americans and Caucasians on the Hopkins Verbal Learning Test–Revised, Brief Visuospatial Memory Test–Revised, Stroop Color and Word Test, and Wisconsin Card Sorting Test 64-Card Version

**Marc A. Norman[1], David J. Moore[1], Michael Taylor[1,3], Donald Franklin Jr.[1],
Lucette Cysique[1,2], Chris Ake[1], Deborah Lazarretto[1], Florin Vaida[1], Robert K. Heaton[1],
and the HNRC Group**

[1]Department of Psychiatry, University of California, San Diego, San Diego, CA, USA
[2]Brain Sciences, University of New South Wales, Sydney, NSW, Australia
[3]Department of Psychology, San Diego State University, San Diego, CA, USA

Memory and executive functioning are two important components of clinical neuropsychological (NP) practice and research. Multiple demographic factors are known to affect performance differentially on most NP tests, but adequate normative corrections, inclusive of race/ethnicity, are not available for many widely used instruments. This study compared demographic contributions for widely used tests of verbal and visual learning and memory (Brief Visual Memory Test–Revised, Hopkins Verbal Memory Test–Revised) and executive functioning (Stroop Color and Word Test, Wisconsin Card Sorting Test–64) in groups of healthy Caucasians ($n = 143$) and African Americans ($n = 103$). Demographic factors of age, education, gender, and race/ethnicity were found to be significant factors on some indices of all four tests. The magnitude of demographic contributions (especially age) was greater for African Americans than for Caucasians on most measures. New, demographically corrected T-score formulas were calculated for each race/ethnicity. The rates of NP impairment using previously published normative standards significantly overestimated NP impairment in African Americans. Utilizing the new demographic corrections developed and presented herein, NP impairment rates were comparable between the two race/ethnicities and were unrelated to the other demographic characteristics (age, education, gender) in either race/ethnicity group. Findings support the need to consider extended demographic contributions to neuropsychological test performance in clinical and research settings.

*Keywords:* Neuropsychology; Executive functioning; Memory; Cognition; African American.

www.manaraa.com

## INTRODUCTION

Learning, memory, and executive functioning are core components of comprehensive neuropsychological (NP) assessment batteries. Accurate classification of NP impairment in these domains is especially important for the differential diagnosis of many neurologic conditions. Unfortunately, some of the most widely used neuropsychological tests do not have available norms that are corrected for race/ethnicity differences, despite research showing that differential ethnicity backgrounds affect NP performance, along with other demographic variables such as age, education, and gender (Heaton, Miller, Taylor, & Grant, 2004). Inadequate normative sampling and standards may lead to neuropsychological misclassification and may particularly contribute to misdiagnosis of African Americans.

Researchers have begun examining demographic influences on learning and memory performance in an effort to produce normative standards among minority groups such as in African Americans—for example, California Verbal Learning Test (CVLT; Norman, Evan, Miller, & Heaton, 2000); and the Third Edition of the Wechsler Memory Scale (WMS–III; Heaton, Taylor, & Manly, 2003—and in Spanish-speaking Hispanics— Hopkins Verbal Learning Test–Revised (HVLT–R) and Brief Visuospatial Memory Test–Revised (BVMT–R; Cherner et al., 2007). However, data on a wider range of neuropsychological tasks are lacking. Regardless of the domain or racial group under study, race/ethnicity typically influences scores on NP measures (Manly, Schupf, Tang, & Stern, 2005).

Several studies have demonstrated lower NP performance among African Americans than among Caucasians on a wide variety of NP measures (Diehr et al., 2003; Diehr, Heaton, Miller, & Grant, 1998; Gladsjo et al., 1999; Heaton et al., 2004; Heaton et al., 2003; Norman et al., 2000; Rilling et al., 2005). Importantly, it has been shown that these differences persist even when groups were matched for other demographic factors, including age, gender, and education, and reading ability to a lesser extent (Manly et al., 2005).

Accurate classification of the level of NP impairment in diverse racial groups has pragmatic clinical relevance to neuropsychologists. Without race/ethnicity-corrected scores in the clinical setting, a substantial number of normal African Americans patients might be incorrectly classified as neuropsychologically impaired and misdiagnosed. For example, Norman et al. (2000) demonstrated that 46% of African Americans were classified as NP impaired (i.e., NP test T-score < 40) on the California Verbal Learning Test (CVLT) Trials 1–5 using the original Delis et al. (Delis, Kramer, Kaplan, & Ober, 1987) norms, which were based upon a predominantly Caucasian standardization sample. When CVLT norms were corrected for age, gender, and education, but not race, 36% of African Americans were still classified as NP impaired. Once race was sufficiently accounted for in the equation, only 17.8% of African Americans were classified as neuropsychologically impaired, demonstrating that the new CVLT norms clearly improved the proportion of individuals scoring greater than one standard deviation below the mean. Misdiagnoses for neurodegenerative disorders or other conditions that affect brain functions have serious implications in terms of public health consequences as well as social and healthcare consequences for the patients and their families. Accurate classification of NP impairment among African Americans is equally important in research settings for similar reasons.

Most existing normative data that are published in test manuals lack information about race/ethnicity influences on test performance. The Hopkins Verbal Learning Test–Revised (HVLT–R) is a widely used task of verbal learning and memory (Brandt & Benedict, 2001), using 12 words belonging to three semantic categories. Six alternate forms facilitate reducing practice effects on repeated administrations. The standardization group's age range was from 15 to 92 ($M = 59.0$, $SD = 18.6$), and education ranged from 2 to 20 years ($M = 13.4$, $SD = 2.9$); 79% were women. The normative sample for the HVLT–R included 1,179 adults; however, racial/ethnicity demographics were not provided. In the HVLT–R manual, stepwise multiple regression examined the effects of age, education, and gender for HVLT–R Total Recall, Delayed Recall, Percent Retained, and Recognition Discrimination. Age accounted for a considerable amount of variance, but education and gender were not found to significantly contribute to test performance. Cherner et al. (2007) contend that a limitation of the original norms was that the reference group was highly educated and had suboptimal representation of low levels of education. Because of this reference group limitation, the rate of NP impairment may be erroneously elevated among lower educated persons.

The Brief Visual Memory Test–Revised (BVMT–R) is a short task of visual memory (Benedict, 1997). As with the HVLT–R, there are six different versions, which allow for repeat testing with reduced practice effects. Similar to the HVLT–R, the manual describes a standardization group of 588 healthy English-speaking adults (171 college students and 471 community-dwelling participants) between the ages of 18 and 79 years ($M = 38.6$, $SD = 18.0$) and with a mean education of 13.4 years ($SD = 1.8$). African Americans accounted for 14.5% of the standardization sample; however, the authors did not provide information concerning whether and how race/ethnicity related to BVMT–R performance. The BVMT–R and HVLT–R produce indices of Total Recall, Delayed Recall, and Percent Retained, and a Recognition Discrimination Index.

The Stroop Color and Word Test consists of speeded trials of Word Reading, Color Naming, and Color–Word Interference. Numerous versions of the Stroop exist, and the version used in the current study assigns a score for each trial based on the number of words read or colors named in 45 seconds (Golden, 1978). The normative sample mentioned in the 2002 manual (Golden & Freshwater, 2002) includes the previous normative group ($n = 100$) from the original manual (Golden, 1978) as well as 300 additional cases collected between 1977

**TABLE 1**
General demographic information on the two groups

| | Caucasian (n = 143) | | | African American (n = 103) | | | |
|---|---|---|---|---|---|---|---|
| | Mean | (SD) | [range] | Mean | (SD) | [range] | p |
| Age | 37.6 | (12.3) | [20–66] | 40.6 | 12.3 | [20–69] | .06 |
| Education | 14.1 | (2.4) | [8–20] | 13.8 | (2.1) | [8–19] | .37 |
| Sex(% female) | 31 | | | 50 | | | .003 |

and 1997 (Golden & Freshwater, 2002). Age and education showed significant associations with Stroop scores, and, as such, the manual includes predicted scores for each trial based on these two demographic characteristics. Gender effects on the Stroop have been examined, but have been found to be inconsistent and confounded by sampling concerns; however, the racial characteristics of the original or total normative samples were not described.

The Wisconsin Card Sorting Test–64-Card Version (WCST–64) is a computerized test of executive function that requires strategic planning and the ability to use environmental feedback to shift cognitive set (Kongs, Thompson, Iverson, & Heaton, 1993). The normative sample for the WCST–64 consisted of 445 adults ages 18–89 years ($M = 49.83$, $SD = 17.92$). Education ranged from 6 to 20 years ($M = 14.95$, $SD = 2.97$), and 23% of the sample was female. Unfortunately, information about race/ethnicity was not routinely collected and therefore was not available for analysis. The manual states that hierarchical polynomial regressions were used to examine the effects of age, gender, and education. Age demonstrated a significant quadratic relationship with WSCT–64 scores and accounted for 1.4% to 18.9% of the variance in scores. Education accounted for 1.3% to 7.7% of the variance in scores after adjusting for age. There were no significant gender effects after accounting for age and education.

This current study was designed to provide improved, demographically corrected normative standards among healthy samples of African Americans and Caucasians on the HVLT–R, BVMT–R, Stroop Color and Word Test, and Wisconsin Card Sorting Test–64-Card Version. The project has two specific aims: (a) to analyze the effects of demographic variables, including race/ethnicity (i.e., African American and Caucasian) on test performance and classification accuracy (normal vs. abnormal); and (b) to develop normative equations that correct for all relevant demographic characteristics (age, education, gender, and race/ethnicity) to provide a more accurate classification of NP performance. We predict that Caucasian versus African American race/ethnicity will significantly contribute to NP performance, and that these differences will support the assertion that verbal and visual learning and memory as well as executive function measures require race/ethnicity corrections in order to correctly categorize NP impairment among African Americans (Manly & Echemendia, 2007).

# METHOD

## Participants

The sample consisted of 246 healthy individuals recruited as comparison participants (HIV-uninfected controls) in a longitudinal study of HIV-infected participants at the University of California, San Diego (UCSD) HIV Neurobehavioral Research Center (HNRC). A total of 143 participants self-identified as Caucasian, and 103 self-identified as African American. Trained research associates used structured interviews and administered screening questionnaires to potential participants to assess inclusion/exclusion criteria prior to study enrollment. Exclusionary criteria for all subjects included any history of neurological disorders, current substance use disorders, and other conditions (e.g., psychiatric disorder with psychotic features, medications with central nervous system, CNS, effects) known to affect neurocognitive performance. The UCSD Human Research Protections Program approved the protocol. Demographic information is presented in Tables 1 and 2. The cross-sectional, stratified sample ranged in age from 20 to 65 years. The two samples (African American and Caucasian) did not differ significantly in terms of age or education, but the Caucasian group contained a smaller proportion of females (31% vs. 50%).

**TABLE 2**
Normative sample cell counts by age and education

| | Age range (years) | | | | |
|---|---|---|---|---|---|
| Education (years) | <30 | 30–39 | 40–49 | 50–59 | 60+ |
| Caucasian | | | | | |
| <10 | 0 | 1 | 1 | 0 | 0 |
| 10–11 | 4 | 2 | 2 | 1 | 0 |
| 12 | 5 | 6 | 5 | 5 | 2 |
| 13–15 | 12 | 4 | 13 | 7 | 0 |
| 16 | 7 | 5 | 4 | 5 | 0 |
| >16 | 2 | 6 | 6 | 6 | 0 |
| African American | | | | | |
| <10 | 0 | 1 | 0 | 0 | 0 |
| 10–11 | 1 | 1 | 4 | 1 | 1 |
| 12 | 6 | 4 | 6 | 3 | 1 |
| 13–15 | 8 | 9 | 10 | 8 | 4 |
| 16 | 4 | 4 | 5 | 1 | 1 |
| >16 | 0 | 3 | 1 | 2 | 2 |

Participants were asked to self-identify their own race/ethnicity, and this identification was used to define the African American and Caucasian groups used in this study. Years of education were determined using a previously defined and standardized procedure where education level ranges from 0–20 based on number of years of schooling completed (Heaton et al., 2004). For example, a high-school graduate receives 12 years of education, and a person with a bachelor's degree receives 16 years of education.

### Neuropsychological assessment

Participants completed an NP test battery of which a subset of two memory and two executive function tests were examined for this study, because these tests lacked race/ethnicity corrections as compared to other tests in the battery. Trained psychometrists following instructions from the respective manuals completed administration and scoring. Analyzed measures included Form A of the Hopkins Verbal Learning Test–Revised (Brandt & Benedict, 2001), Display A from the Brief Visual Memory Test–Revised (Benedict, 1997), the Stroop Color–Word Interference Test (Golden, 1978), and the Wisconsin Card Sorting Test–64 Computer Version (Heaton, Chelune, Talley, Kay, & Curtiss, 1993). We evaluated Total Recall across three learning trials and Delayed Recall for the HVLT–R and BVMT–R. Additionally, total numbers of correct items identified with the 45-s trials were analyzed for Stroop Word Reading, Color Naming, and Color–Word formats. For the WCST, scores analyzed included Total Errors, Perseverative Errors, and Conceptual Level Responses.

### Data analysis

The distributions of all scores were examined. Although distributions of test raw scores were non-normal, parametric statistics were confirmed with nonparametric versions of the same statistical comparisons, and tails of distributions were similar between racial groups assuming symmetry in impairment rates. Effect sizes were measured with the unbiased Cohen's $d$ (Hedges & Olkin, 1985). This study was powered to detect a small effect size.

In the first step, African American and Caucasian group scores were compared to analyze the effects of race/ethnicity on test performance. Next, linear regression was used to examine the effects of age, education, and gender; this was done separately for African American and Caucasian groups because it was determined that they had somewhat different age effects and (to a lesser degree) education effects. Partial regressions were then run to examine the independent contribution of age, education, and gender on measures in each group (Caucasian and African American).

HVLT–R, BVMT–R, Stroop, and WCST–64 raw scores for the total subject group were converted into quantiles and were mapped into the corresponding quantiles of a standard normal distribution. These scores were then converted into normalized scaled scores with a mean of 10 and standard deviation of 3. We used a subset of individuals ($n = 208$) from the present study and some additional normal subjects from other ethnicities to create a census-matched subset of individuals to generate the scaled scores as described below, but results from other ethnicities were not used in subsequent analyses that focused on African Americans and Caucasians. The rationale for adding these additional individuals for raw-score to scaled-score conversions was to reflect in the scaled scores the major ethnic group composition reported in the 2000 US census. These individuals met the same screening procedures as those for the study population. The resulting census-matched proportions of race/ethnicity categories were 68.7% Caucasian, 13.5% African American, 13.0% Hispanic, and 4.8% other race/ethnicities. Scaled-score conversion tables for all variables are presented in Tables 3–5.

In the next step, fractional polynomial multiple regression was employed to develop demographically corrected prediction equations on the Caucasian and African American samples (respective $n$s = 143 and 103) for each NP test scaled score using the methods outlined by Royston and Altman (1994; also see Cherner et al., 2007; Heaton et al., 2004). Separate regressions were run for each race/ethnicity, and the predictors included age, education, and sex. The fractional polynomial method developed by Royston and Altman uses an interactive algorithm to evaluate the influence of combinations of predictors with predetermined exponents (–2, –1, –0.5, 0, 0.5, 1, 2, 3; the coefficient of 0 stands for the natural logarithm transformation). The algorithm compares all

**TABLE 3**
Raw- to scaled-score conversions for the BVMT–R and HVLT–R

| | Raw | | | |
| | BVMT–R | | HVLT–R | |
| Scaled | Total Recall | Delayed Recall | Total Recall | Delayed Recall |
|---|---|---|---|---|
| 17 | 36 | | 36 | |
| 16 | | | | |
| 15 | 34 – 35 | | 35 | |
| 14 | 33 | 12 | 34 | |
| 13 | 32 | | 33 | 12 |
| 12 | 30 – 31 | | 31 – 32 | |
| 11 | 28 – 29 | 11 | 30 | 11 |
| 10 | 26 – 27 | | 29 | |
| 9 | 24 – 25 | 10 | 27 – 28 | 10 |
| 8 | 21 – 23 | 9 | 26 | |
| 7 | 19 – 20 | 8 | 24 – 25 | 9 |
| 6 | 16 – 18 | 7 | 22 – 23 | 8 |
| 5 | 14 – 15 | 5 – 6 | 21 | 7 |
| 4 | 10 – 13 | 4 | 20 | 5 – 6 |
| 3 | 0 – 9 | 3 | 16 – 19 | |
| 2 | | 0 – 2 | 0 – 15 | 4 |
| 1 | | | | 0 – 3 |

*Note.* BVMT–R = Brief Visuospatial Memory Test–Revised. HVLT–R = Hopkins Verbal Learning Test–Revised.

**TABLE 4**
Raw- to scaled-score conversions for the Stroop Color and Word Test

| | Raw | | |
| | Stroop Color and Word Test | | |
| Scaled | Word Reading | Color Naming | Color–Word |
|---|---|---|---|
| 18 | ≥145 | ≥107 | |
| 17 | 134–144 | 100–106 | ≥65 |
| 16 | 128–133 | 97–99 | 63–64 |
| 15 | 123–127 | 93–96 | 59–62 |
| 14 | 118–122 | 89–92 | 56–58 |
| 13 | 114–117 | 85–88 | 53–55 |
| 12 | 109–113 | 80–84 | 49–52 |
| 11 | 106–108 | 76–79 | 46–48 |
| 10 | 101–105 | 74–75 | 42–45 |
| 9 | 97–100 | 70–73 | 39–41 |
| 8 | 89–96 | 66–69 | 36–38 |
| 7 | 83–88 | 62–65 | 32–35 |
| 6 | 77–82 | 58–61 | 29–31 |
| 5 | 71–76 | 49–57 | 25–28 |
| 4 | 67–70 | 43–48 | 22–24 |
| 3 | 66 | 40–42 | 0–21 |
| 2 | <66 | 0–39 | |

**TABLE 5**
Raw- to scaled-score conversions for the WCST–64

| | Raw | | |
| Scaled | Total Errors | Perseverative Errors | Conceptual Level Responses |
|---|---|---|---|
| 18 | | 0–3 | |
| 17 | 0–6 | | ≥58 |
| 16 | 7 | | 57 |
| 15 | 8 | | 56 |
| 14 | | 4 | 54–55 |
| 13 | 9–10 | | 53 |
| 12 | 11 | 5 | 51–52 |
| 11 | 12 | 6 | 49–50 |
| 10 | 13–15 | 7 | 45–48 |
| 9 | 16–19 | 8 | 39–44 |
| 8 | 20–22 | 9–10 | 34–38 |
| 7 | 23–28 | 11–13 | 28–33 |
| 6 | 29–32 | 14–15 | 20–27 |
| 5 | 33–35 | 16–18 | 16–19 |
| 4 | 36–39 | 19–26 | 13–15 |
| 3 | 40–48 | 26–41 | 6–12 |
| 2 | ≥49 | ≥42 | ≤5 |
| 1 | | | |

*Note.* WCST–64 = Wisconsin Card Sorting Test 64-Card Version.

sets of predictors using these transformations to generate the final optimal fit. The residuals from the optimal regression equations were converted to T-scores with a mean of 50 and a standard deviation of 10. As designed, the resultant T-scores are not correlated with age, sex, or education for either racial group.

## RESULTS

### Neuropsychological test performance

In the first step, African American and Caucasian raw scores on each of the neuropsychological measures examined in this study were compared to analyze the

effects of race/ethnicity. Table 6 demonstrates significant Caucasian and African American differences on all measures, such that Caucasians performed better in each instance. Table 6 also depicts medium to large effect sizes on most learning, memory and executive functioning indices; the only exceptions were small to medium effect sizes on HVLT–Delayed Recall and Stroop Color Naming and Word Reading.

### Regression analysis

For each of the 10 test scores, stepwise linear regressions were then conducted separately for each group (African

**TABLE 6**
Neuropsychological test performance in Caucasians and African Americans

| | Caucasian (n = 143) | African American (n = 103) | | |
| | M (SD) | M (SD) | p | Cohen's d |
|---|---|---|---|---|
| BVMT–R Total Recall | 26.5 (5.9) | 22.7 (6.8) | .0003 | −0.60 |
| BVMT–R Delayed Recall | 10.2 (1.7) | 8.7 (2.4) | <.0001 | −0.74 |
| HVLT–R Total Recall | 29.2 (3.9) | 26.8 (4.9) | .0002 | −0.55 |
| HVLT–R Delayed Recall | 10.4 (1.9) | 9.4 (2.3) | .0016 | −0.48 |
| Stroop Word Reading | 101.9 (14.4) | 96.2 (16.9) | .007 | −0.37 |
| Stroop Color Naming | 76.4 (10.8) | 70.8 (13.0) | .0008 | −0.47 |
| Stroop Color–Word | 45.0 (9.5) | 38.2 (10.2) | <.0001 | −0.69 |
| WCST–64 Total Errors | 15.6 (7.8) | 22.1 (10.2) | <.0001 | 0.73 |
| WCST–64 Perseverative Errors | 7.6 (3.9) | 11.0 (4.2) | .0002 | 0.63 |
| WCST–64 Conceptual Level Responses | 44.3 (11.3) | 35.5 (14.3) | <.0001 | −0.69 |

*Note.* Test performance is shown as raw scores. BVMT–R = Brief Visuospatial Memory Test–Revised. HVLT–R = Hopkins Verbal Learning Test–Revised. WCST–64 = Wisconsin Card Sorting Test 64-Card Version.

**TABLE 7**
Full model $R^2$ and partial $R^2$ for the effect of age, education, and gender in each group on memory measures

| | Caucasian (n = 143) | | | African American (n = 103) | | |
|---|---|---|---|---|---|---|
| | $R^2$ | Partial $R^2$ | 95% CI | $R^2$ | Partial $R^2$ | 95% CI |
| BVMT–R Total Recall | .11** | | .02, .20 | .28*** | | .14, .42 |
| Age | | .05** | .00, .12 | | .21*** | .07, .35 |
| Education | | .05** | .00, .12 | | .05** | .00, .13 |
| Sex | | .02 | .00, .06 | | .03* | .00, .09 |
| BVMT–R Delayed Recall | .06* | | .00, .13 | .29*** | | .15, .43 |
| Age | | .03* | .00, .08 | | .24*** | .10, .38 |
| Education | | .04* | .00, .10 | | .03* | .00, .09 |
| Sex | | .00 | .00, .02 | | .03* | .00, .09 |
| HVLT–R Total Recall | .19*** | | .07, .30 | .15*** | | .03, .27 |
| Age | | .00 | .00, .02 | | .07** | .00, .16 |
| Education | | .18*** | .06, .29 | | .05* | .00, .13 |
| Sex | | .03* | .00, .08 | | .03* | .00, .09 |
| HVLT–R Delayed Recall | .18*** | | .06, .29 | .14*** | | .02, .26 |
| Age | | .01 | .00, .04 | | .05* | .00, .13 |
| Education | | .14*** | .04, .24 | | .08** | .00, .18 |
| Sex | | .05** | .00, .12 | | .01 | .00, .05 |

*Note.* BVMT–R = Brief Visuospatial Memory Test–Revised. HVLT–R = Hopkins Verbal Learning Test–Revised. WCST–64 = Wisconsin Card Sorting Test 64-Card Version. *$p \leq$ .05. **$p \leq$ .01. ***$p \leq$ .001.

American and Caucasian) to determine the proportion of variance accounted for by age, education, and gender (Tables 7 and 8). None of the fractional polynomials were significant predictors.

### Memory

Table 7 shows information related to the demographic influences on learning and memory performance in Caucasians and African Americans independently. When considering the partial $R^2$ results, only the African American group showed a significant effect of age, and this was true for all measures (especially robust for BVMT–R measures). Total demographic effects ($R^2$s) were higher for African Americans due to greater age effects on the BVMT–R, whereas more comparable effects were seen for the HVLT–R. Although education was a significant independent predictor of memory test performance for all measures in both groups, the education effects on the verbal (HVLT–R) measures were especially robust for the Caucasian group. Gender effects were absent or modest for both groups on most measures, with women performing better, and there were no systematic differences for the Caucasians versus African Americans.

### Executive functioning

Comparable results for Stroop and WCST–64 measures are presented in Table 8. As was the case for measures of visual learning and memory (BVMT–R), only the African American groups showed very large

independent effects of age on all of the Stroop indices (Word Reading, Color Naming, and Color–Word). Only the African American group also showed significant gender effects on Stroop Color–Word (Interference condition) and Color Naming, with women performing faster. On the WCST–64 measures, both race/ethnicity groups demonstrated medium-sized age effects (typically somewhat larger for Caucasians), and usually small to medium education effects. Neither racial/ethnicity group showed gender effect on this test.

### Normative T-score derivation

As described in the Method section, fractional polynomial regression analyses were conducted to derive normative scores that would correct for the observed demographic effects on normal test performance. This procedure began with the conversion of raw scores to normalized scaled scores (mean = 10, $SD = 3$) on all test measures (see Tables 3–5 for these conversions).

To examine the diagnostic ("normal" vs. "abnormal") classification accuracy of the new T-score conversions with more complete demographic corrections, we compared the impairment rates in both samples with those using previously published normative data (Benedict, 1997; Brandt & Benedict, 2001; Golden & Freshwater, 2002; Kongs et al., 1993) that did not correct for race/ethnicity. The formulas used to generate the results for the new T-scores are included in the Appendix. Subjects were considered impaired if their T-score was less than 40 (Heaton et al., 2004; Taylor & Heaton, 2001).

**TABLE 8**
Full model $R^2$ and partial $R^2$ for the effect of age, education, and gender in each group on executive measures

| Measures | Caucasian (n = 143) | | | African American (n = 103) | | |
|---|---|---|---|---|---|---|
| | $R^2$ | Partial $R^2$ | 95% CI | $R^2$ | Partial $R^2$ | 95% CI |
| *Stroop measures* | | | | | | |
| Word Reading | .01 | | .00, .04 | .20*** | | 07, .33 |
|   Age | | .00 | .00, .02 | | .12*** | .00, .24 |
|   Education | | .00 | .00, .02 | | .06* | .00, .15 |
|   Sex | | .00 | .00, .02 | | .03 | .00, .09 |
| Color Naming | .01 | | .00, .04 | .26*** | | .12, 40 |
|   Age | | .00 | .00, .02 | | .16*** | .03,.29 |
|   Education | | .01 | .00, .04 | | .04* | .00, .11 |
|   Sex | | .00 | .00, .02 | | .05** | .00, .13 |
| Color–Word | .15*** | | .04, .25 | .39*** | | .25, .53 |
|   Age | | .11*** | .01, .20 | | .28*** | .14 42 |
|   Education | | .04* | .00, .10 | | .03* | .00, .09 |
|   Sex | | .00 | .00, .02 | | .10*** | .00, .21 |
| *WCST–64 measures* | | | | | | |
| Total Errors | .22*** | | .10, .34 | .19** | | .06, .32 |
|   Age | | .17*** | .06, .28 | | .13** | .01, .25 |
|   Education | | .07*** | .00, .15 | | .06** | .00, .15 |
|   Sex | | .00 | .00, .02 | | .00 | .00, .03 |
| Perseverative Errors | .20*** | | .09, .31 | .14** | | .02, .26 |
|   Age | | .17*** | .06, .28 | | .09** | .00, .19 |
|   Education | | .05** | .00, .12 | | .05* | .00, .13 |
|   Sex | | .00 | .00, .02 | | .00 | .00, .03 |
| Conceptual Level Responses | .20*** | | .09, .31 | .20*** | | .07, .33 |
|   Age | | .15*** | .04, .26 | | .13*** | .01, .25 |
|   Education | | .07*** | .00, .15 | | .07** | .00, .16 |
|   Sex | | .00 | .00, .02 | | .00 | .00, .03 |

*Note.* WCST–64 = Wisconsin Card Sorting Test 64-Card Version. *$p \leq$ .05. **$p \leq$ .01. ***$p \leq$ .001.

Figure 1 shows the results for the African American group. When applying previously published normative corrections to this sample, 24–49% of normal individuals were classified as NP impaired depending on the test score examined. Using our newly generated normative data, the impairment rates significantly improved and ranged from 13–16%. The impairment rates for the African American sample with the previously published norms are significantly greater than what would be expected from the normal distribution with the selected cutoff of 1 standard deviation (Golden, 1978). All comparisons of impairment rates among African Americans using previously published normative corrections as compared to the newly generated normative corrections were statistically significant with the exception of the HVLT–R Delayed Recall measure, which approached significance ($p = .08$).

The new normative correction formulas improved the consistency of impairment rates across test scores for the Caucasian sample as well (see Figure 2). Impairment rates for these norms ranged from 12 to 17% as compared to 8 to 26% using previously published normative data. The newly developed WCST–64 norms produced impairment rates more aligned with the expected impairment rates and were significantly lower than impairment rates with previously published norms on all WCST–64 indices.

**DISCUSSION**

This study complements previous literature on demographic corrections for neuropsychological test norms by examining a broader range of memory and executive functioning measures and specifically examining the effect of African American versus Caucasian race/ethnicity on test performance. These findings strongly support the use of separate norms for African American and Caucasian examinees on the tests used here and, when combined with previously published results in the same ability domains, on learning, memory, and executive functioning measures more generally. Consistent with prior findings on the Wechsler Intelligence and Memory Scales (Heaton et al., 2003) and expanded Halstead–Reitan Battery (Heaton et al., 2004), we found, in our sample of 103 African Americans and 143 Caucasians, that African American participants obtained lower raw scores on visual and verbal learning and memory and executive functioning measures.
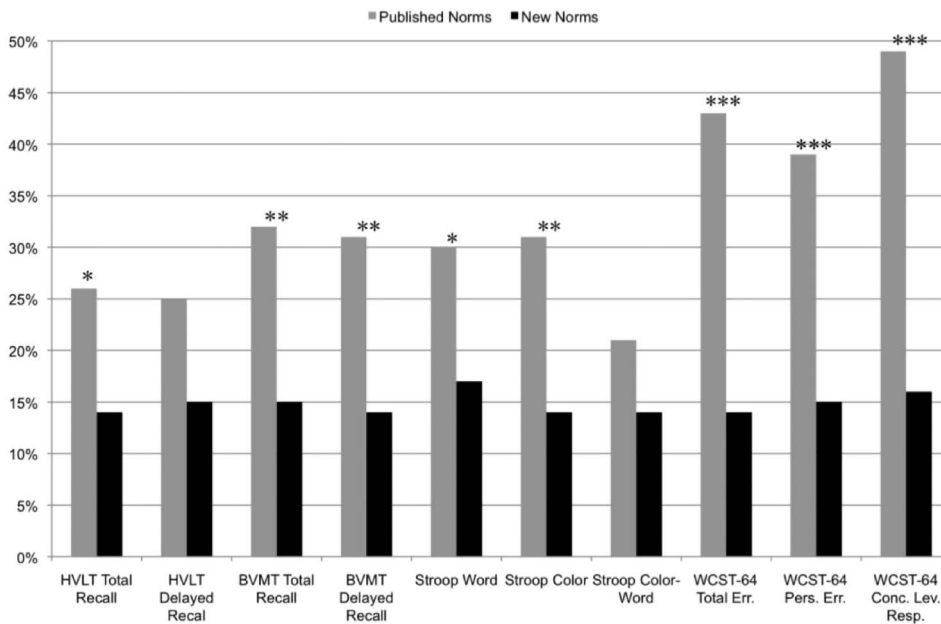
**Figure 1.** Percentage of normal African American sample classified as "impaired" (1-*SD* cutoff) by published norms versus new, demographically corrected norms. HVLT = Hopkins Verbal Learning Test. BVMT = Brief Visuospatial Memory Test. WCST–64 = Wisconsin Card Sorting Test 64-Card Version. *$p \leq .05$. **$p \leq .01$. ***$p \leq .001$.
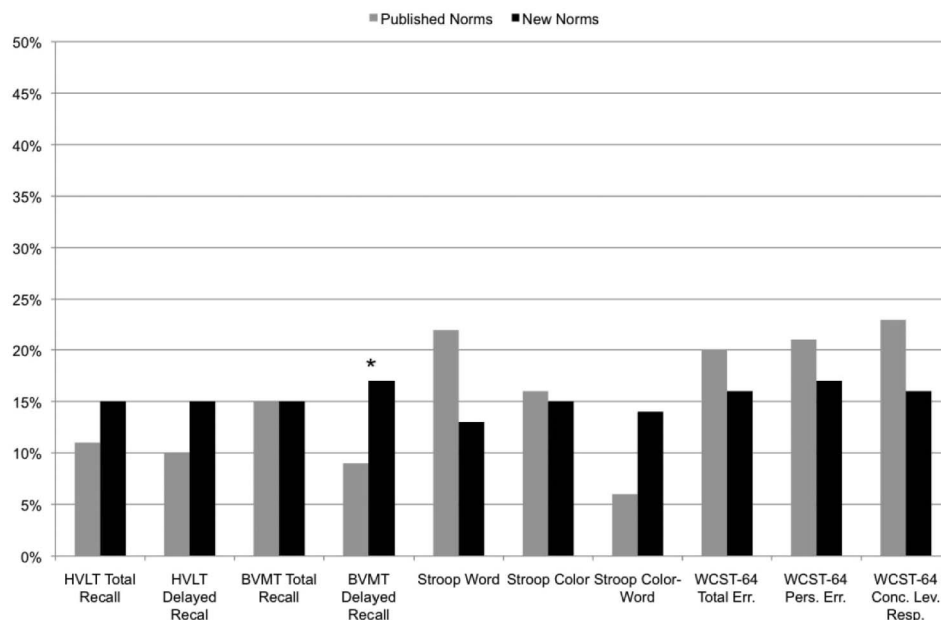


**Figure 2.** Percentage of normal Caucasian sample classified as "impaired" (1-*SD* cutoff) by published norms versus new, demographically corrected norms. HVLT = Hopkins Verbal Learning Test. BVMT = Brief Visuospatial Memory Test. WCST–64 = Wisconsin Card Sorting Test 64-Card Version. *$p \leq .05$. **$p \leq .01$. ***$p \leq .001$.

There are multiple background differences between African American and Caucasian adults within U.S. society today that may place African Americans at a disadvantage on standardized NP testing. The observed raw NP score differences may be consistent with disparities in quality of formal and informal educational experiences; however, other factors may also contribute to these discrepancies. It is considered unlikely that race has a direct causal effect on differences in adult cognition, so race/ethnicity is viewed as a proxy for other factors, much like what has been discussed about education (Manly, Byrd, Touradji, & Stern, 2004). Factors potentially contributing to raw NP score differences between African American and Caucasian groups may include academic exposure, education quality, academic resources, acculturation, socioeconomic status, social exposure, "test wiseness," societal discrimination (Byrd, Sanchez, & Manly, 2005; Manly et al., 2004), and lifelong experiences contributing to low group and self-expectations (Steele & Aronson, 1995).

There are few opportunities in the literature to compare our raw score results with those reported by other investigators. Whereas this study found about a 2-point (raw score) difference on HVLT–R Total Recall performances between Caucasians and African Americans, Morgan et al. (Morgan, Marsiske, & Whitfield, 2008) found a 4-point difference and less variability. The current study demonstrated moderate to large race/ethnicity effect sizes, but the raw score differences between Caucasians and African Americans do not seem to be especially large (e.g., an average of only 1.5 points on BVMT–R Delayed Recall is associated with a medium to large effect size; see Table 6). However, these differences were sufficiently robust to cause unacceptably large "impairment" classification rates in the African American sample (Figure 1).

Although concerns might be raised that the method of raced-based norming could "overcorrect" performances of neurologically impaired African Americans (making them less sensitive to disease), this could be said as well for norms that correct for older age, lower education levels, or any demographic characteristic that is associated with lower test performance in normal people. In our view, the most important function of norms is to maintain an acceptable and *consistent* level of diagnostic specificity (accuracy in classifying normal people as normal) for people regardless of their demographic characteristics. Our data suggest that the norms presented here result in rates of impairment that are comparable and are within statistical expectations for a healthy population, for both our Caucasian and African American participants.

As addressed by Byrd et al. (2005) and others, the term "race" is an arbitrary distinction and difficult to operationalize. Often race is based on skin color and self-identification. As Gasquoine (2009) notes, race is a social definition rather than a scientific classification, and race is not homogeneous. Devising ways to understand the influences of ethnicity and race on NP tests will become increasingly complex as rates of self-identified multiracial individuals rise.

Given the unclear relationship of "race" on cognition, some suggest recording, quantifying, and modeling the effects of all background factors that can influence cognitive development and test performance. Gasquoine (2009) and others have advocated that an alternative approach to race/ethnicity-based norms is to estimate preexisting neuropsychological status based on a case-by-case basis from regular normative tables. On the other hand, Gasquoine acknowledges that there is little empirical support for this technique, and there is no agreed-upon method for establishing NP status on a case-by-case basis. Furthermore, accurate retrospective collection of such complex data across the lifespan is very difficult (Byrd et al., 2005).

Also, a subjective interpretation of cognitive deficit will most likely have the effect of wide variations in the impairment classifications of minorities between different clinical neuropsychologists. Instead, the use of the more general race/ethnicity proxy (with all its shortcomings) in normative corrections should at least enhance consistency/reliability and may greatly

reduce the probability of incorrectly attributing cognitive and possibly CNS abnormalities to normal African Americans.

Of course, clinical interpretation of neuropsychological data should not strictly rely upon use of norms, but should also consider the appropriateness of available norms in relation to each person's background, including social, educational, and medical history, and other factors (i.e., psychiatric, substance use, etc.). In particular, diagnostic sensitivity and specificity are likely to vary when norms are applied to people whose backgrounds differ significantly from those represented in the normative sample populations (Heaton et al., 2004).

In addition, it is important to note that demographically corrected norms are intended to reflect the *difference* between current performance and a best estimate of the person's expected "normal" performance (i.e., in the absence of CNS abnormality). Such norms are less appropriate, at best, when the goal is to determine the person's *absolute* level of ability (e.g., in relation to requirements of specific everyday tasks and activities).

Following the derivation of separate normative equations and confirming adequate normative distributions, we found that our new demographically corrected formulas provided significantly improved impairment estimates. These data suggest that scores that have not been corrected for race/ethnicity classify 31–32% of the African American sample with visual learning and memory impairment, 25–26% as having verbal learning and memory impairment, and 24–49% as having executive dysfunction. These percentages are substantially higher than expected values in any normative population. In contrast, when the African American scores were corrected for race/ethnicity, the average impairment frequencies dropped to expected levels. The overestimation of impairment with existing normative data can lead to misclassification and/or misdiagnosis of African American individuals and can have serious negative consequences for the patients and their families. Misdiagnosis and misclassification are problematic in clinical, forensic, and research applications of neuropsychology; however, few NP norms account for these demographic variables.

The present study also demonstrated that the demographic contributions of age, education, and gender to NP test performances were somewhat different for African Americans as compared to Caucasians. The contribution of age tended to be stronger for African American participants on the BVMT–R, HVLT–R, and Stroop tests, but less pronounced on the WCST–64. The current study was not designed to explore why demographic factors exhibit stronger influences among African Americans, although large age effects for African Americans as compared to Caucasians have been observed in other large U.S. samples and on other neuropsychological tests (Heaton et al., 2004). Because these differential effects of demographics are not well understood, they require additional careful investigation (especially taking into account age-related medical conditions and associated treatments that could differ across ethnicity groups).

The current study is limited, as are others of the same type, in terms of failing to provide insights into the factors that contribute to racial differences on these memory and executive function measures. As discussed earlier, Manly et al. (1998) and others suggest that educational quality, exposure, and other factors might play a role in the poorer observed performance of African Americans on these neuropsychological tests. The amount of education may be less important than the quality of one's education, as measured by reading scores. Dotson et al. (Dotson, Kitner-Triolo, Evans, & Zonderman, 2008) and Manly et al. (Manly et al., 1999; Manly, Jacobs, Touradji, Small, & Stern, 2002) found that literacy was a better predication of cognitive scores than education. In an African American sample, Dotson and colleagues (2008) used memory, naming, fluency, visuospatial, attention, and psychomotor scores and regressed them on sex, age, literacy, and education scores. They did not find a unique contribution of education after literacy was added to the model; however, this study only included African American participants. The present study did not measure literacy, but, as with previous studies, education was found to be a significant predictor to cognitive scores. The measurement of these factors remains elusive, however, as effects of educational opportunities and importance within the cultural experience, and other potentially important factors are complex and difficult to determine retrospectively (e.g., asking an adult about parental influences and early school experiences; Byrd et al., 2005). On the other hand, current attempts at understanding these factors are starting to emerge, and multifactorial models involving psychological factors, stress factors, and social and cognitive factors have been proposed (Mays, Cochran, & Barnes, 2007). An additional complexity is that it is likely that some or all of the factors influencing NP test performance have changed over generations and continue to do so. For example, it is likely that the educational quality for 30-year-old and 60-year-old African Americans has been quite different (probably more so for than for Caucasians in the US).

An additional limitation in this research is the ambiguity in classifying race or ethnicity. While "race" and "ethnicity" are often interchangeably used in this area of research, they are not equivalent terms. Given that there are no biological race/ethnicity markers, group identification has been pragmatically based on self-identification—and this is the approach that was used on the current study. Race is more than just skin color, and there may be multiple ethnic groups within a race. Some argue that the inability to specifically identify and characterize race/ethnicity should preclude demographic corrections; however, even with this limitation, the current data demonstrate excessive rates of diagnostic error if clinicians use norms that are not corrected for race/ethnicity. In particular, our findings with the new T-score conversions suggest greater and more equal specificity, with regard to race/ethnicity, within the healthy population than that achieved by the published norms. Although we have no data concerning sensitivity of the new norms to CNS compromise, sensitivity also is likely to be more equivalent among demographic groups (e.g.,

Heaton, Ryan, & Grant, 2009). Despite limitations, we believe that the current quantitative standards provide a substantial improvement for the classification of neurocognitive impairment status in self-identified African Americans.

Finally, it is important to acknowledge that our current sample size was relatively small, and we were unable to cross-validate the normative distribution with an independent sample. We recommend caution when using these normative data with individuals over age 60 or with other groups not well represented in our normative sample. There were relatively few individuals with fewer than 10 years of education enrolled in this study, and therefore caution should be used when applying these normative corrections to persons with such low levels of education. In addition, all participants in this study were from the San Diego area, and participants were carefully screened to exclude anyone with neuromedical or developmental histories suggesting any increased risk for CNS compromise. As such, generalizability of these results and associated normative standards to other, ostensibly normal, African American and Caucasian groups cannot be assumed. To partially address this question, we applied the demographically corrected norms in the WAIS–III/WMS–III/WIAT–II Scoring Assistant program (Heaton et al., 2003; Psychological Corporation, 1999; WAIS–III = Wechsler Adult Intelligence Scale–Third Edition; WMS–III = Wechsler Memory Scale–Third Edition; WIAT–II = Wechsler Individual Achievement Test–Second Edition) to the current samples' results on three WAIS–III subtests (Letter–Number Sequencing, Digit Symbol Coding, and Symbol Search). These latter norms were based upon a large, national standardization sample from all U.S. regions and correct for all demographic variables that were examined in the current study (age, education, gender, and African American vs. Caucasian race/ethnicity). We reasoned that application of these norms to the current samples' WAIS–III results would provide some indication of their representativeness of the much larger national sample. Ideally, the mean (SD) T-scores would approach 50 (10) and would not differ for the two race/ethnicity groups in the study.

For Letter–Number Sequencing, the mean (SD) T-scores were 53.0 (9.8) for our African American Group and 51.8 (9.4) for our Caucasian group (p-value for group difference = .35). On the WAIS–III Processing Speed Index (which combines Digit Symbol and Symbol Search), the respective scores were 54.6 (10.7) for our African American group and 52.6 (9.9) for our Caucasian group (p = .14). The fact that both of our race/ethnicity groups performed slightly better than the national standardization samples on these WAIS–III tests may reflect our (arguably) more stringent neuromedical screening procedures and/or slight regional differences. Also, however, these results indicate that, relative to normal expectations for African American and Caucasians in the U.S., our race/ethnicity groups performed comparably. This suggests that our groups' findings on the memory and executive function tests are unlikely to be overestimating the race/ethnicity bias in the previously published norms.

Our results for the HVLT–R and BVMT–R are limited to Form A of these measures, and future studies will focus on assessing the need for specific corrections for all the multiple forms of these measures. Additionally, it is important to assess whether or not the demographic corrections can be validated in a clinical sample, showing equivalent results across the various demographic categories.

## REFERENCES

Benedict, R. (1997). *Brief Visuospatial Memory Test–Revised*. Odessa, FL: Psychological Assessment Resources.

Brandt, J., & Benedict, R. (2001). *Hopkins Verbal Learning Test–Revised*. Lutz, FL: Psychological Assessment Resources.

Byrd, D. A., Sanchez, D., & Manly, J. J. (2005). Neuropsychological test performance among Caribbean-born and U.S.-born African American elderly: The role of age, education and reading level. *Journal of Clinical and Experimental Neuropsychology*, 27(8), 1056–1069.

Cherner, M., Suarez, P., Lazzaretto, D., Fortuny, L. A., Mindt, M. R., Dawes, S., et al. (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. *Archives of Clinical Neuropsychology*, 22(3), 343–353.

Delis, D. C., Kramer, J. H., Kaplan, E., & Ober, B. A. (1987). *California Verbal Learning Test–Manual*. San Antonio, TX: The Psychological Corporation.

Diehr, M. C., Cherner, M., Wolfson, T. J., Miller, S. W., Grant, I., & Heaton, R. K. (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. *Journal of Clinical and Experimental Neuropsychology*, 25(4), 571–585.

Diehr, M. C., Heaton, R. K., Miller, W., & Grant, I. (1998). The Paced Auditory Serial Addition Task (PASAT): Norms for age, education, and ethnicity. *Assessment*, 5(4), 375–387.

Dotson, V. M., Kitner-Triolo, M., Evans, M. K., & Zonderman, A. B. (2008). Literacy-based normative data for low socioeconomic status African Americans. *The Clinical Neuropsychologist*, 22(6), 989–1017.

Gasquoine, P. G. (2009). Race-norming of neuropsychological tests. *Neuropsychology Review*, 19(2), 250–262.

Gladsjo, J. A., Schuman, C. C., Evans, J. D., Peavy, G. M., Miller, S. W., & Heaton, R. K. (1999). Norms for letter and category fluency: Demographic corrections for age, education, and ethnicity. *Assessment*, 6(2), 147–178.

Golden, C. J. (1978). *Stroop Color and Word Test: A manual for clinical and experimental uses*. Chicago, IL: Stoelting.

Golden, C. J., & Freshwater, S. M. (2002). *Stroop Color and Word Test: A manual for clinical and experimental uses*. Lutz, FL: Psychological Assessment Resources.

Heaton, R. K., Chelune, G. J., Talley, J. L., Kay, G. G., & Curtiss, G. (1993). *Wisconsin Card Sorting Test manual*. Odessa, FL: Psychological Assessment Resources.

Heaton, R. K., Miller, S. W., Taylor, M. J., & 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, R. K., Ryan, L., & Grant, I. (2009). Demographic influences and use of demographically corrected norms in neuropsychologial assessment. In I. Grant & K. M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric and neuromedical disorders* (pp. 127–155). New York, NY: Oxford University Press.

Heaton, R. K., Taylor, M. J., & Manly, J. J. (2003). Demographic effects and use of demographically corrected norms for the WAIS–III and WMS–III. In D. Tulsky, D. Saklofske, G. Chlune, R. Heaton, R. Ivnik, R. Bornstein, A. Prifitera, & M. Ledbetter (Eds.), *Clinical interpretation of the WAIS–III and WMS–III (practical resources for the mental health professional)* (pp. 183–210). San Diego, CA: Academic Press.

Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. Orlando, FL: Academic Press.

Kongs, S. K., Thompson, L. L., Iverson, G. L., & Heaton, R. K. (1993). *Wisconsin Card Sorting Test–64 card version professional manual*. Odessa, FL: Psychological Assessment Resources.

Manly, J. J., Byrd, D. A., Touradji, P., & Stern, Y. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology*, 11(1), 37–46.

Manly, J. J., & Echemendia, R. J. (2007). Race-specific norms: Using the model of hypertension to understand issues of race, culture, and education in neuropsychology. *Archives of Clinical Neuropsychology*, 22(3), 319–325.

Manly, J. J., Jacobs, D. M., Sano, M., Bell, K., Merchant, C. A., Small, S. A., et al. (1999). Effect of literacy on neuropsychological test performance in nondemented, education-matched elders. *Journal of the International Neuropsychological Society*, 5(3), 191–202.

Manly, J. J., Jacobs, D. M., Touradji, P., Small, S. A., & Stern, Y. (2002). Reading level attenuates differences in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society*, 8(3), 341–348.

Manly, J. J., Miller, S. W., Heaton, R. K., Byrd, D., Reilly, J., Velasquez, R. J., et al. (1998). The effect of African-American acculturation on neuropsychological test performance in normal and HIV-positive individuals [The HIV Neurobehavioral Research Center (HNRC) Group]. *Journal of the International Neuropsychological Society*, 4(3), 291–302.

Manly, J. J., Schupf, N., Tang, M. X., & Stern, Y. (2005). Cognitive decline and literacy among ethnically diverse elders. *Journal of Geriatric Psychiatry and Neurology*, 18(4), 213–217.

Mays, V. M., Cochran, S. D., & Barnes, N. W. (2007). Race, race-based discrimination, and health outcomes among African Americans. *Annual Review of Psychology*, 58, 201–225.

Morgan, A. A., Marsiske, M., & Whitfield, K. E. (2008). Characterizing and explaining differences in cognitive test performance between African American and European American older adults. *Experimental Aging Research*, 34(1), 80–100.

Norman, M. A., Evan, J. D., Miller, S. W., & Heaton, R. K. (2000). Demographically corrected norms for the California Verbal Learning Test. *Journal of Clinical and Experimental Neuropsychology*, 22(1), 80–94.

Psychological Corporation. (1999). *WAIS–III/WMS–III/WIAT–II scoring assistant*. San Antonio, TX: Author.

Rilling, L. M., Lucas, J. A., Ivnik, R. J., Smith, G. E., Willis, F. B., Ferman, T. J., et al. (2005). Mayo's Older African American Normative Studies: Norms for the Mattis Dementia Rating Scale. *The Clinical Neuropsychologist*, 19(2), 229–242.

Royston, P., & Altman, G. G. (1994). Regression using fractional polynomials of continuous covariates: Parsimonious parametric modeling. *Applied Statistics*, 43, 429–467.

Steele, C. M., & Aronson, J. (1995). Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology*, 69(5), 797–811.

Taylor, M. J., & Heaton, R. K. (2001). Sensitivity and specificity of WAIS–III/WMS–III demographically corrected factor scores in neuropsychological assessment. *Journal of the International Neuropsychological Society*, 7(7), 867–874.

**APPENDIX**

<div align="center">

**Table A1**
Normative formulas for Caucasians and African Americans

</div>

| | *T-score formulas* |
|---|---|
| ***Caucasian*** | |
| BVMT Total Recall | [(Total learning scaled score – (0.2589 × (edu – 14.11) + (–0.0515) × (age – 37.62) + 0.9276 × sex + 10.0712))/2.8912] × 10 + 50 |
| BVMT Delayed Recall | [(Delayed recall scaled score – (0.2084 × (edu – 14.11) + (–0.0286)× (age – 37.62) + 10.3007))/2.6989] × 10 + 50 |
| HVLT Total Recall | [(Total learning scaled score – (0.5225 × (edu – 14.11) + 1.0035 × gender + 10.1738))/2.5927] × 10 + 50 |
| HVLT Delayed Recall | [(Delayed recall scaled score – (0.5414 × (edu – 14.11) + 1.7324 × gender + 9.9285))/2.6989] × 10 + 50 |
| Stroop Word Reading | [(Word reading scaled score – (0.0819× (edu – 13.92) + 0.0038 × (age – 36.17) + (–0.4022) × gender + 10.2102) – (–0.00003))/2.9435] × 10 + 50 |
| Stroop Color Naming | [(Color naming scaled score – (0.0941 × (edu – 13.92) + 10.4444))/2.8101] × 10 + 50 |
| Stroop Color–Word | [(Color–word scaled score – (0.2479 × (edu – 13.92) + (–0.0828) × (age – 36.22) + 10.3968))/2.6002] × 10 + 50 |
| Stroop Interference | [(Interference scaled score – (0.2346 × (edu – 13.92) + (–0.0762) × (age – 36.22) + 0.7465 × sex + 9.9952))/2.6342] × 10 + 50 |
| WCST–64 Total Errors | [(Total errors scaled score – (0.3187× (edu – 14.13) + (–0.01) × (age – 37.37) + 0.1608 × gender + 10.4049) – (–0.0017))/2.674] × 10 + 50 |
| WCST–64 Perseverative Errors | [(Perseverative errors scaled score – (0.2357 × (edu – 14.13) + (–0.0941) × (age – 37.37) + 0.0341 × gender + 10.33) – (–0.0012))/2.5506] × 10 + 50 |
| WCST–64 Conceptual Level Responses | [(Conceptual level responses scaled score – (0.3223× (edu – 14.13) + (–0.0941) × (age – 37.37) + 0.0577 × gender + 10.4292) – (–0.0016))/2.568] × 10 + 50 |
| ***African American*** | |
| BVMT Total Recall | [(Total learning scaled score – (0.2834 × (edu – 13.86) + (–0.1125) × (age – 40.63) + 1.0394 × sex + 8.0679))/2.5701] × 10 + 50 |
| BVMT Delayed Recall | [(Delayed recall scaled score – (0.2267 × (edu – 13.86) + (–0.12.62) × (age – 40.63) + 0.8593 × sex + 7.691))/2.5197] × 10 + 50 |
| HVLT Total Recall | [(Total learning scaled score – (0.2917 × (edu – 13.86) + (–0.0644) × (age – 40.63) + 1.1462 × sex + 8.3063))/2.8333] × 10 + 50 |
| HVLT Delayed Recall | [(Delayed recall scaled score – (0.3986 × (edu – 13.86) + (–0.0733) × (age – 40.63) + .9145 × sex + 8.2753))/3.1354] × 10 + 50 |
| Stroop Word Reading | [(Word reading scaled score – (0.3557 × (edu – 13.92) + (–0.0866× (age – 40.67) + 1.2315 × sex + 8.3263) – (0.00095))/2.8127] × 10 + 50 |
| Stroop Color Naming | [(Color naming scaled score – (0.3102 × (edu – 13.94) + (–0.1006) × (age – 40.67) + 1.4915 × sex + 8.2672))/2.6643] × 10 + 50 |
| Stroop Color–Word | [(Color–word scaled score – (0.2363 × (edu – 13.94) + (–0.1219) × (age – 40.67) + 1.9479 × sex + 7.449))/2.2658] × 10 + 50 |
| Stroop Interference | [(Interference scaled score – ((–0.0303) × (age – 40.67) + 1.4688 × sex + 8.1343))/2.663] × 10 + 50 |
| WCST–64 Total Errors | [(Total errors scaled score – (0.3321× (edu – 13.97) + (–0.0838) × (age – 40.7) + 0.3215× sex + 8.1621) – (–0.0006))/2.7911] × 10 + 50 |
| WCST–64 Perseverative Errors | [(Perseverative errors scaled score – (0.3599 × (edu – 13.97) + (–0.0776) × (age – 40.7) + (–0.1093) × sex + 8.5524) – (0.0006))/3.0124] × 10 + 50 |
| WCST–64 Conceptual Level Responses | [(Conceptual level responses scaled score – (0.4002× (edu – 13.97) + (–0.0874) × (age – 40.7) + 0.2546 × gender + 8.2248) – (0.0006))/2.5887] × 10 + 50 |

*Note.* BVMT = Brief Visuospatial Memory Test. HVLT = Hopkins Verbal Learning Test. WCST–64 = Wisconsin Card Sorting Test 64-Card Version. Sex: male = 0; female = 1. Edu = education: Years of education were determined using a previously defined and standardized procedure where education level ranges from 1–20 based on number of years of schooling completed (Heaton et al., 2004).