Education correction using years in school or reading grade-level equivalent? Comparing the accuracy of two methods in diagnosing HIV-associated neurocognitive impairment

Mona Rohit[1], **Andrew Levine**[1], **Charles Hinkin**[2,3], **Shogik Abramyan**[1], **Ernestine Saxton**[1,4], **Miguel Valdes-Sueiras**[1], and **Elyse Singer**[1]

[1]Department of Neurology, National Neurological AIDS Bank—University of California, Los Angeles, California

[2]Department of Psychiatry & Biobehavioral Sciences, David Geffen School of Medicine—University of California, Los Angeles, California

[3]Department of Psychiatry, Veterans Administration Greater Los Angeles Health Care System, Los Angeles, California

[4]Department of Neurology, Charles R. Drew University of Medicine and Science, Los Angeles, California

Abstract

Neuropsychological tests generally require adjustments for years of education when determining the presence of neurocognitive impairment. However, evidence indicates that educational quality, as assessed with reading tests, may be a better reflection of educational attainment among African Americans. Thus, African Americans with poor educational quality may be incorrectly classified with neurocognitive impairment based on neuropsychological tests. We compared the accuracy of neuropsychological test scores standardized using reading grade-equivalent *versus* years of education in predicting neurocognitive impairment among a sample of Whites and African-American adults who were HIV+. Participants were examined by a neurologist and classified with or without HIV-associated neurocognitive disorders according to accepted criteria. Participants were also classified as impaired *versus* not impaired based on their neuropsychological test scores standardized by 1) self-reported education or 2) WRAT-3 reading grade-level. Cross tabulation tables were used to determine agreement of the two methods in detecting impairment. Among African-Americans, standardized scores derived from reading scores had greater specificity than those derived from years of education (84.1% *vs.* 77.3). Among the Whites, correction based on years of education had both greater specificity and sensitivity. The results suggest that reading tests may be a useful alternative for determining NCI among African Americans.

Keywords

HIV; Dementia; Neuropsychology; WRAT-3; Education; Reading

Correspondence and reprint requests to: Andrew Levine, Ph.D., National Neurological AIDS Bank, 11645 Wilshire Blvd., Suite 770, Los Angeles, CA 90025. ajlevine@mednet.ucla.edu.

Introduction

Neuropsychological (NP) tests are frequently used to diagnose neurocognitive impairment (NCI) either individually or in conjunction with other methods, such as neurological examination. These tests use actuarial methods to compare an individual's performance to that of the average individual with similar demographic characteristics. Most NP tests use normative data that has been stratified based on age and education level, as these factors have consistently been found to be related to cognitive ability. Scoring based on such stratified systems is more accurate than if one were to group all participants together, as abilities change across the life-span and individuals with higher levels of education tend to have an advantage on most NP tests. However, this latter point is based on the assumption that quality of education is similar across schools, and that there is little variability in ability among individuals who attain the same level of education. It is well-established that quality of education is highly variable across individuals, reflecting such factors as different schools, teaching methods, teacher quality, pupil/teacher ratios, presence of special facilities, length of school year, attendance, and peer characteristics (Manly et al., 2002). Thus, while two individuals may each report 12 years of education, they may have vastly different quality of education based on these factors. This could result in significant discrepancies in their standardized scores based on education stratification that considers only years of education. The individual with a poorer quality of education may, when compared to the average individual with a similar education level, appear more impaired than s/he actually is. Conversely, the individual with the higher quality of education may still fall within the normal range of performance despite having a true decline in ability. Furthermore, differences in NP scores are greater in those with lower education levels as opposed to individuals with higher education levels (Ostrosky-Solis et al., 1998), perhaps reflecting the greater variability in educational quality in pre-college schooling. Therefore, mounting evidence suggests that length of schooling does not necessarily equal quality. Stratifying normative data based on the former may decrease accuracy in assigning standardized scores, and therefore diagnosis of impairment.

One common method for determining quality of education is the reading test. Reading ability is highly correlated with direct measures of quality of education (e.g., teacher/student ratios, teacher education) and academic achievement (Wilkinson, 1993). Johnstone et al. (1997), in suggesting that there is an inherent weakness in estimating premorbid abilities based on education because this assumes that individuals matched for years of education will perform at the same level of cognitive functioning, proposed that an individual's reading grade-level is a more accurate assessment of cognitive abilities and premorbid intelligence. Additionally, it is well-established that reading ability (albeit not reading comprehension) is relatively stable in the presence of brain dysfunction (Christensen et al., 1991; Crawford et al., 1992; Klesges & Sanchez, 1981; Klesges & Troster, 1987), with the exception of focal brain damage to areas subserving reading.

Certain ethnic groups, because of socioeconomic disparity, may be more likely to be misclassified based on normative data stratified by years of education. Numerous studies have indicated that the quality of education received by African-Americans, particularly those who are economically disadvantaged, may vary markedly from the education received by Whites. Performance on NP tests of reading (Boekamp et al., 1995), naming (Lichtenberg et al., 1994; Roberts & Hamsher, 1984; Ross et al., 1995), and nonverbal abilities (Adams et al., 1982; Anger et al., 1997; Bernard, 1989; Brown et al., 1991; Campbell et al., 1996; Heverly et al., 1986; Miller et al., 1993) have consistently shown that African Americans score lower than Whites on verbal and nonverbal cognitive tasks while accounting for the same socioeconomic status and stated years of education. According to Manly et al. (2002), these findings largely reflect the differences in quality of education received by African Americans compared to Whites. This discrepancy between stated years of education and quality of

education has been demonstrated repeatedly (Baker et al., 1996; O'Bryant et al., 2005; Ryan et al., 2005; Wilson et al., 2003). For example, O'Bryant and colleagues (2005) recently reported that among a sample of psychiatric patients, a significantly larger discrepancy between self-reported education and a reading grade-equivalent based on the reading subtest of the Wide Range Achievement Test (WRAT) was found among the African American as compared to the Whites in their sample. If we are to assume that reading tests such as the WRAT are valid indicators of education quality, then this suggests that educational stratification based on years of education may not be the best method for some ethnic groups.

One implication from these observations is that African Americans are more likely than Whites to be misdiagnosed as neurocognitively impaired when measures based on stated years of education are used (Klusman et al., 1991; Manly et al., 1998; Stern et al., 1992; Welsh et al., 1995). This is of particular concern with regards to the human immunodeficiency virus (HIV), to which African Americans are disproportionately affected (CDC, 2003). HIV often leads to NCI, which includes minor cognitive/motor disorder (MCMD) and HIV-associated dementia (HAD). These conditions are usually diagnosed based largely upon the NP examination. Thus, due to the reasons discussed above, groups such as African Americans may be more prone to misclassification of NCI. Recently, Ryan et al. (2005) presented evidence that supports this hypothesis. With the aim of investigating the effects of education quality on NP performance, Ryan et al. examined a sample of 200 urban-dwelling individuals, 51% of whom were African American and 24% of whom were Hispanic. The authors hypothesized that discrepancy between years of education and reading grade-level, but not ethnicity, would account for differences in NP performances among a cohort of African American, Hispanic, and White adults who were HIV+. Their findings were consistent with this, and they also found that the minority participants had larger discrepancies between their reported years of education and reading grade-level. Further, they found that when reading grade-level was substituted for education when obtaining norm-based standardized scores, impairment rates fell considerably for all participants, and somewhat more so for the minority groups. Thus, although their study lacked a criterion against which to assess the diagnostic accuracy of the reading grade-level method, the results indicate that African Americans may be more prone to be misclassified with NCI.

In the current study, we continue this line of inquiry by examining discrepancies in NP test scores that are standardized based on self-reported years of education *versus* WRAT reading grade-equivalent in a sample of White and African American individuals who are HIV+. In addition, we examined which of the methods is more accurate in assigning neurocognitive diagnoses, with a neurologist's diagnosis (naïve of NP test results) used as the criterion variable. It has been observed within our clinic that the examining neurologist often judges the patient to be less impaired based on their examination than does the neuropsychologist, who assigns diagnoses based on scores collected via psychometric tests. We hypothesize that education level established by the WRAT would result in higher NP scores and more accurate classification (i.e., better specificity) of neurocognitive status among our African American participants.

Method

Participants

The sample consisted of 113 HIV+, English-speaking adults who were participants of the National Neurological AIDS Bank (NNAB) study. Of the 113 participants, 62 were non-Hispanic African Americans and 51 non-Hispanic Whites. Hispanic individuals were not included in the current analysis because of the high percentage within our clinic that are monolingual Spanish-speaking or whose English fluency is limited. A total of 18 females were included, comprising 15.9% of the sample. Mean age was 42.7 years (*sd* = 8.5). Mean self-

reported years of education was 13 years (*sd* = 2.3). Demographic characteristics by ethnic group are presented in Table 1.

Procedure

As a part of the NNAB, participants received comprehensive physical and neurological examinations at study entry and again at regular intervals of 6 to 12 months. HIV status was determined via HIV ELISA and confirmed with western blot and0or HIV PCR. Individuals with history of seizure disorder, learning disability, head injury resulting in loss of consciousness lasting more than 1 hour, or opportunistic infections affecting the central nervous system (e.g., toxoplasmosis, progressive multifocal leukoencephalopathy, and cryptococcal meningitis) were excluded from the analyses.

Neuropsychological measures—Upon study entry, all participants were administered a comprehensive battery of NP tests to help determine the presence of NCI. This was performed by psychometrists trained by a board-certified neuropsychologist (C.H.). The battery consisted of the 1) Trail Making Test (TMT), Forms A & B (Army Individual Test Battery, 1944), 2) Grooved Pegboard (Klove, 1963), 3) Symbol Search, a subtest of the WAIS-III battery (Wechsler, 1997), 4) Digit Symbol, also a subtest of the WAIS-III, 5) Paced Auditory Serial Addition Test (PASAT) (Gronwall, 1974), and 6) Controlled Oral Word Association Test (COWAT) (Benton & Hamsher, 1989). Standardized T-scores were derived from published normative data (Heaton, 1991; Wechsler, 1997) that did not stratify for ethnicity.

Reading test—The reading subtest from the Wide Range Achievement Test, Third Edition, or WRAT-3 (Wilkinson, 1993), was used to obtain grade-level estimates based on reading ability. Raw scores were converted to standardized scores using normative data provided in the test manual. Grade-level equivalents ranged from pre-kindergarten through 8th grade, as well as 'high school' and 'post high school.' Consistent with studies similar to ours (Ryan et al., 2005), those classified as having 'high school' reading equivalents were assigned a grade level of 12 and those with 'post high school' equivalents were assigned a grade level of 13.

Neurological diagnosis—All participants were examined by a board-certified study neurologist upon study entry and prior to neuropsychological testing. The examination included observation of the participant, cognitive screening, blood and CSF analysis, and a standard neurological examination. Following the neurological examination, the study neurologist entered a preliminary neurocognitive diagnosis based upon their findings. Ninety-six of the 113 participants were classified according to American Academy of Neurology criteria (1991) as one of the following: 1) neurocognitively normal: patients with no evidence of impairment based on cognitive screening; 2) subsyndromic: demonstrating subtle deficits not meeting criteria for HIV-related neurocognitive disorder (i.e., evidence of cognitive abnormalities which do not impair the subject's ability to carry out activities of daily living and do not manifest as clinical symptoms); 3) possible or probable MCMD: mildly impaired cognitive ability, reported symptoms of cognitive decline and disability, and diagnostic evaluation revealing another possible cause for impairment (i.e. possible MCMD) or ruling out other cause of impairment (i.e., probable MCMD); 4) possible or probable HAD: meets criteria for dementia, reports symptoms of cognitive decline and significant disability, and diagnostic evaluation revealing another possible cause for impairment (i.e. possible HAD) or ruling out other cause of impairment (i.e., probable HAD). Individuals believed to have neurocognitive impairment due to other causes (e.g. head injury, opportunistic infection, neoplasm) were excluded from the analyses. The percentage of NCI cases among African Americans and Whites was 48% and 29%, respectively.

Data analyses—Raw scores were transformed into standardized *T*-scores. Published normative data for all measures are stratified by age. For the purposes of this study, we standardized the raw scores in two ways: 1) with self-reported years of education and 2) using the grade-level obtained via the WRAT-3 Reading test in place of education. Thus, two sets of *T*-scores were derived for comparison in our analyses. Four sets of analyses were then performed. First, we characterized the cohort with regard to demographic and virologic variables, as well as discrepancy between self-reported education and reading grade-level. Because of our interest in understanding the contribution of ethnicity to this discrepancy, ANOVAs were used to compare our two groups (White and African American) across these variables. Second, in order to examine the agreement between the two correction methods, paired-sample *t*-tests were performed for each of the NP measures within each of the two groups. Third, between-group comparisons (White *vs.* African American) on the NP measures based on the two correction methods were done using 2×2 (group × method) mixed-model ANOVA in order to investigate interactions between these factors. The fourth and final set of analyses was aimed at examining which of the correction methods results in greater diagnostic accuracy. This was accomplished by examining agreement of NCI (as determined via NP measures) and neurological diagnosis of HIV-related neurocognitive disorder (as determined by the neurologist) among the 96 participants who were formally diagnosed. For the former, classification of NCI was defined as an average *T*-score of less than 40, based on all NP measures. For the latter, those diagnosed as neurocognitively normal or subsyndromic were combined into a single group labeled *neurocognitively asymptomatic*, while those diagnosed with possible or probable MCMD or HAD were grouped together as *neurocognitively symptomatic*. Rates of agreement for impairment (i.e., sensitivity) and no impairment (i.e., specificity) were determined for the entire cohort, and then separately for each ethnic group, using cross tabulation frequency tables.

Results

Table 1 displays the results of demographic and reading level comparisons between the two ethnic groups. The groups did not differ with regards to age, self-reported education, absolute CD4+ cell count, or plasma HIV viral load. The African American group had a significantly greater proportion of females ($\chi^2 = 7.01$, $p = .01$). African Americans also had a lower reading grade-level (mean = 8.6, *sd* = 4.4) as compared to Whites (mean = 11.4, *sd* = 3) ($F[1,111] = 15.02$, $p < .001$), but not a lower level of self-reported education ($F[1,111] = .63$, $p = .67$).

In order to rule out gender as a confounding factor for between group comparisons, NP test *T*-scores were compared between males and females across the entire sample. For the education-based norms, gender differences were found only for the grooved pegboard dominant ($F[1,111] = 6.65$, $p = .01$) and nondominant [$F(1,111) = 7.34$, $p = .01$], with females' time being significantly slower. For the reading-based scores, both dominant [$F(1,111) = 9.35$, $p = .003$] and non-dominant [$F(1,111) = 10.36$, $p = .002$] grooved pegboard were again slower for females, and they also had fewer words on the COWAT [$F(1,111) = 4.3$, $p = .04$]. Gender was therefore entered as a covariate in a later ANOVA comparing the two groups on these measures.

Within the White sample, paired-sample *t*-tests revealed that WRAT-corrected scores resulted in significantly higher *T*-scores (see Table 2). This was true for the PASAT ($t = -3.53$, $p = .001$), COWAT ($t = -3.16$, $p = .003$), Trails Making Test Parts A ($t = 22.33$, $p = .02$) & B ($t = -2.4$, $p = .02$), and Grooved Pegboard-dominant hand ($t = -2.62$, $p = .01$). For the African American sample, when education-corrected scores were used, participant's scores were higher for Symbol Search ($t = 3.56$, $p = .001$), PASAT ($t = 4.48$, $p < .001$), COWAT ($t = 3.97$, $p < .001$), and Letter-Number Sequencing ($t = 3.15$, $p = .003$) (see Table 3). Conversely, WRAT-corrected norms produced higher scores for Trails Making Test, Parts A ($t = -7.14$, $p < .001$)

and B ($t = -7.01$, $p < .001$), as well as Grooved Pegboard-dominant hand ($t = -5.03$, $p < .001$) among the African Americans.

Mixed-model ANOVA was used to determine the individual and interactive influence of normative method and ethnicity on test performance. Interaction effects were found for Trail Making Test-Part A ($F = 6.78$, $p = .01$) and Part B ($F = 6.65$, $p = .01$), as well as Grooved Pegboard-dominant hand after co-varying for gender ($F = 4.09$, $p = .04$). In each instance, the scores among African American participants increased more than Whites when the WRAT-3 grade level was used to correct scores. No main or interaction effects were found for the nondominant hand. The opposite trend was found among other measures, including the PASAT ($F = 28.14$, $p < .001$), Letter-Number Sequencing ($F = 6.32$, $p = .01$), and Symbol Search ($F = 5.77$, $p = .02$). On these measures, it was the White participants whose scores increased with the WRAT-3, while scores of African Americans tended to decrease on average. Gender was also entered as a covariate for the COWAT, for which a similar interaction was found ($F = 21.38$, $p < .001$). Symbol did not have an interaction effect, but did have a main effect for correction method ($F = 4.67$, $p = .03$), with WRAT-3 resulting in higher scores.

Finally, diagnostic accuracy rates were determined first for the entire sample and then for each group. For the entire sample, sensitivity (i.e., impairment according to both NP testing and neurologist's examination) was somewhat better when self-reported grade level was used to norm test scores (Table 4). Particularly, sensitivity was 46.2% for WRAT reading grade-level and 55.8% for self-reported education. Conversely, specificity (i.e., rating of unimpaired by both NP testing and neurologist) was somewhat better when WRAT reading grade-level was used (84.1%) as compared to grade attainment (77.3%). Examination of the individual ethnic cohorts revealed interesting differences. While WRAT and grade attainment resulted in similar accuracy rates among the African American cohort (59%), the former allowed greater specificity (77.8% *vs.* 55.6%) but lower sensitivity (48.4% *vs.* 61.3%) (see Table 5). Conversely, among the White cohort, overall accuracy was slightly better when grade attainment was used as a correction factor (72% *vs.* 68.1%). Further, both sensitivity and specificity were greater when this method was used (47.6% *vs.* 42.9% and 92.3% *vs.* 88.5%, respectively) (see Table 6).

Discussion

In standardizing NP test scores, education level is frequently considered in order to correct for the effects that schooling has on cognitive ability. However, previous studies have found that grade attainment is not the best indicator of educational quality and may result in underestimating NP test performances among African Americans, thereby making them more prone to an erroneous diagnosis of NCI. As a result, reading level has been suggested to be a more accurate reflection of one's true educational quality, especially among African Americans.

Despite similar self-reported years of education among African Americans and Whites, reading grade-levels based on the WRAT-3 proved to be significantly lower for the former group. This is consistent with past studies (Manly et al., 1998; Manly et al., 2002), which demonstrated that African-Americans had attained a lower quality of education (operationalized as WRAT-3 reading grade-equivalent) than Whites matched for years of education.

Within each ethnic group, there were significant differences between NP test scores using the two correction methods. Among the African-Americans, significantly higher scores were obtained on visual attention and psychomotor tests (e.g., Trail Making Test and Grooved Pegboard) when WRAT-3 correction was used. Conversely, this group achieved higher scores primarily on measures of executive functioning and verbal attention (e.g., Letter/Number

Sequencing, COWAT, and PASAT) when scores were corrected using years of education. Among the Whites, scores obtained via WRAT-3 correction were consistently higher than those based on grade attainment. However, examination of the differences in scores obtained from the two methods shows most of them to be quite small, on the order of about 1 T-score unit. Thus, while statistically significant, it is unclear that these differences were of clinical significance. The one exception was the PASAT, a measure of working memory. In addition, a significant interaction was seen between ethnicity and correction method across almost all tests, further indicating that NP test scores differed depending on both correction method and ethnicity.

These findings differ somewhat from those of Ryan et al. (2005), who found that the African Americans within their sample, similar to the Whites, had consistently higher scores across tests when the WRAT-3 grade level was used as a correction factor. One explanation may be the significant difference in the actual grade attainment between their White and African American cohorts (14.3 *vs.* 11.7, respectively). In our sample the two groups had equivalent grade attainment. Thus, there was a greater discrepancy between grade attainment and reading grade-equivalent among their African American cohort. In addition, Ryan et al. (2005) examined impairment rates (defined as 1.5 *SD*s below the mean for each measure of interest) only in those individuals who had a significant discrepancy between reported grade attainment and reading grade-equivalent. Therefore, the greater discrepancy between self-reported education and reading grade-level likely resulted in higher scores when the latter was used as a correction method. The disparate findings may also be the result of minor differences among the test batteries used. Ours included primarily measures requiring processing speed, but little in the way of language and reasoning skills, which some might argue are more highly correlated with education. However, previous findings from our group indicate that performance on even simple reaction time measures are predicted by education level (Levine et al., 2004). The findings among our African American sample are perhaps more consistent with those of Johnstone et al. (1997), who found varying rates of NCI depending on whether reading level or years of education was used as a correction factor. Based on a sample of primarily White adults age 40 and under with a history of traumatic brain injury, the authors reported that reading-based scores (derived from the WRAT-R or WRAT-3) were associated with greater impairment on both parts of the Trail Making Test. In addition, this method resulted in a larger discrepancy in scores between cognitive and motor tasks. In contrast, they found that scores based on years of education were associated with greater rates of impairment on motor tasks and nonverbal IQ, but that the discrepancy between cognitive and motor performance was not as remarkable. Thus, the authors suggested that WRAT-based correction is associated with greater variability of impairment across abilities, and that this is perhaps more reflective of the greater sensitivity of this method. Note that Johnstone et al. (1997) did not use the grade-equivalent from the WRAT, but rather calculated *z*-scores derived from the reading subtest score. Z-scores for each of their cognitive domains of interest were then subtracted from the reading z-scores in estimating rates of impairment. Thus, the difference in methodology between their study and ours makes comparison difficult.

Perhaps most striking is the finding that the two correction methods are associated with differential accuracy depending on ethnicity. Although Ryan et al. (2005) found that using reading grade-level (via WRAT-3) as a proxy for years of education lowered rates of impairment (defined as a deviation from the sample mean) across a variety of NP tests, our study is the first to compare the diagnostic accuracy of the two methods using an external criterion (i.e. neurologist's diagnosis). For our entire sample, there was little difference in accuracy rates between the two correction methods, although WRAT-correction led to better specificity while scores based on years of education led to greater sensitivity. More compelling, however, are the findings that the two methods had differential diagnostic accuracy among the two ethnic groups. Consistent with our hypothesis, WRAT-corrected scores were found to

increase specificity rates by over 20% above that of grade attainment-corrected scores (77.8% *vs.* 55.6%) among the African American cohort. Thus, these results support the notion that NP scores derived from self-reported years of education may lead to artificially inflated rates of impairment among this group. However, using WRAT-corrected scores may also have drawbacks, as the sensitivity associated with this method was significantly lower than that of the traditional technique (48.4% *vs.* 61.3%). Among the White cohort, overall accuracy was slightly better when using years of education as the correction factor (68.1% *vs.* 72%). Both sensitivity and specificity decreased by approximately 4% when WRAT-correction was used. Thus, the traditional method appears to be more accurate for Whites. These findings suggest that different correction methods may be appropriate for these two groups. The decision to employ reading grade-level as a correction factor for African Americans will rest upon the tradeoff between sensitivity and specificity.

There were a number of limitations to the current study, which should be considered. First, the WRAT-3 reading test scores were skewed such that most participant scores were in the upper part of the range, suggesting a ceiling effect for this test among our sample. As education level progresses to the high school years, the effect of educational quality is no longer as robust as when comparing participants among lower educational levels (Ostrosky-Solis et al., 1998). This lack of variability can also adversely impact the statistical analysis. Thus, analyzing a sample that has more variability with regards to reading ability will be useful. Second, the study consisted predominantly of males, with females comprising 15.9% of our study population. However, it is worthy to note that men have higher rates of HIV, with women accounting for 22% of HIV infected individuals (CDC, 2003). Therefore, this gender disparity generally reflects the demographics of HIV within the Los Angeles area, where the most common risk behavior for HIV remains male-to-male sexual contact. There was a significant difference in gender between racial groups. Although our analyses co-varied for gender to eliminate possible gender differences, a sample that has a more similar gender composite may be helpful in seeing the effects these normative methods have on diagnoses. Third, the results are based on the premise that reading ability is fundamentally similar among ethnic groups. However, an alternative explanation may be that the WRAT-3 is not appropriate for estimating reading level among African Americans. For example, the words used on the WRAT-3 may be less commonly used within schools that serve primarily African Americans, or within their homes and social settings. Thus, their poorer scores on the WRAT-3 may have been due to lack of familiarity rather than poor educational quality. This fundamental question will require further investigation. Another issue regarding the WRAT-3 is that there may have been greater variance in the abilities of African-Americans lumped in the "high school" reading level as compared to the Whites in that same category among the original WRAT sample. Unfortunately, the WRAT-3 has an inherent weakness in not assigning specific reading grade levels. The norms the WRAT-3 is based on aggregates all subjects with a reading level from ninth to twelfth grade as "high school" and all subjects with a reading level above the twelfth grade as "post high school." Since the time of our study a fourth edition of the WRAT has been published, which has added grade-based norms, thus increasing the utility of the test in differentiating the grade levels within high school. Consequently, employing the WRAT-4 in future studies similar to ours will be of value. Finally, when determining impairment based on NP tests, we employed a cutoff score of 40, which is one standard deviation below the mean of the normative sample to which our cohort was compared. It is possible that this cutoff was not the most appropriate threshold for our sample. Adjusting the cutoff may have resulted in an improvement in overall accuracy rates for both methods examined. In addition, weighting certain tests over others may have increased our accuracy rates. However, while these psychometric issues are highly relevant to the current study, we believe that the current findings are just an initial step towards creating more fitting normative methods for African Americans, and minorities in general. Future studies will likely shed light on the additional psychometric issues that have arisen here.

These results have important implications on the HIV+ population. With the advent and use of highly active antiretroviral therapy (HAART), HIV-infected individuals are living longer and experiencing lower rates of opportunistic infections. However, we are seeing a rising prevalence of other HIV-associated conditions, including neurocognitive disorders (Fischer-Smith & Rappaport, 2005). Moreover, HIV is affecting a growing number of African Americans and other minority populations. Taking this into consideration, it is necessary to have enhanced diagnostic tools with normative data that are more representative of the typical HIV + demographic. These preliminary findings suggest that specifying the most appropriate normative method for individuals from particular backgrounds may significantly reduce misdiagnosis.

Acknowledgments

References

Adams RL, Boake C, Crain C. Bias in a neuropsychological test classification related to education, age, and ethnicity. Journal of Consulting and Clinical Psychology 1982;50:143–145. [PubMed: 7056909]

American Academy of Neurology. Nomenclature and research case definitions for neurologic manifestations of human immunodeficiency virus-type 1 (HIV-1) infection. Report of a Working Group of the American Academy of Neurology AIDS Task Force. Neurology 1991;41:778–785. [PubMed: 2046917]

Anger WK, Sizemore OJ, Grossman SJ, Glasser JA, Letz R, Bowler R. Human neurobehavioral research methods: Impact of subject variables. Environmental Research 1997;73:18–41. [PubMed: 9311528]

Army Individual Test Battery. Manual of Directions and Scoring. Washington, D.C.: War Department, Adjutant General's Office; 1944.

Baker FM, Johnson JT, Velli SA, Wiley C. Congruence between education and reading levels of older persons. Psychiatry Services 1996;47:194–196.

Benton, AL.; Hamsher, KdeS. Multilingual aphasia examination. Iowa City, IA: AJA Associates; 1989.

Bernard LC. Halstead-Reitan Neuropsychological Test performance of black, Hispanic, and white young adult males from poor academic backgrounds. Archives of Clinical Neuropsychology 1989;4:267–274. [PubMed: 14589609]

Boekamp JR, Strauss ME, Adams N. Estimating premorbid intelligence in African-American and white elderly veterans using the American version of the National Adult Reading Test. Journal of Clinical and Experimental Neuropsychology 1995;17:645–653. [PubMed: 8557806]

Brown A, Campbell A, Wood D, Hastings A, Lewis-Jack O, Dennis G. Neuropsychological studies of Blacks with cerebrovascular disorders: A preliminary investigation. Journal of the National Medical Association 1991;83:213–217.

Campbell, A.; Rorie, K.; Dennis, G.; Wood, D.; Combs, S.; Hearn, L. Neuropsychological assessment of African Americans: Conceptual and methodological considerations. Vol. 2. Oakland, CA: Cobb & Henry; 1996.

CDC. HIV/AIDS Surveillance Report, 2003. Vol. 15. Atlanta, GA: US Department of Health and Human Services, Center for Disease Control and Prevention; 2003.

Christensen H, Hadzi-Pavlovic D, Jacomb P. The psychometric differentiation of dementia from normal aging: A meta-analysis. Psychological Assessment 1991;3:147–155.

Crawford JR, Moore JW, Cameron IM. Verbal fluency: A NART-based equation for the estimation of premorbid performance. British Journal of Clinical Psychology 1992;31:327–329. [PubMed: 1393161]

Fischer-Smith T, Rappaport J. Evolving paradigms in the pathogenesis of HIV-1-associated dementia. Expert Reviews in Molecular Medicine 2005;7:1–26. [PubMed: 16321172]

Gronwall, DH. The psychological effects of concussion. Auckland: Oxford University Press; 1974.

Heaton, RK. Comprehensive norms for an expanded Halstead-Reitan Battery: Demographic corrections, research findings, and clinical applications. Odessa, FL: Psychological Assessment Resources, Inc.; 1991.

Heverly LL, Isaac W, Hynd GW. Neurodevelopmental and racial differences in tactile-visual (cross-modal) discrimination in normal black and white children. Archives of Clinical Neuropsychology 1986;1:139–145. [PubMed: 14589647]

Johnstone B, Slaughter J, Schopp L, McAllister JA, Schwake C, Luebbering A. Determining neuropsychological impairment using estimates of premorbid intelligence: Comparing methods based on level of education versus reading scores. Archives of Clinical Neuropsychology 1997;12:591–601. [PubMed: 14590671]

Klesges RC, Sanchez VC. Cross-validation of an index of premorbid intellectual functioning in children. Journal of Consulting and Clinical Psychology 1981;49:141. [PubMed: 7217471]

Klesges RC, Troster AI. A review of premorbid indices of intellectual and neuropsychological functioning: What have we learned in the past five years? The International Journal of Clinical Neuropsychology 1987;9:1–11.

Klove H. Clinical Neuropsychology. The Medical Clinics of North America 1963;47:1647–1658. [PubMed: 14078168]

Klusman LE, Moulton JM, Hornbostel LK, Picano JJ, Beattie MT. Neuropsychological abnormalities in asymptomatic HIV seropositive military personnel. Journal of Neuropsychiatry and Clinical Neuroscience 1991;3:422–428.

Levine A, Miller E, Becker J, Selnes O, Cohen B. Normative data for determining significance of test-retest differences on eight common neuropsychological instruments. The Clinical Neuropsychologist 2004;18:373–384. [PubMed: 15739809]

Lichtenberg PA, Ross T, Christensen B. Preliminary normative data on the Boston Naming Test for an older urban population. The Clinical Neuropsychologist 1994;8:109–111.

Manly JJ, Jacobs DM, Touradji P, Small SA, Stern Y. Reading level attenuates differences in neuropsychological test performance between African American and White elders. The Journal of the International Neuropsychological Society 2002;8:341–348.

Manly JJ, Miller SW, Heaton RK, Byrd D, Reilly J, Velasquez RJ. The effect of African-American acculturation on neuropsychological test performance in normal and HIV-positive individuals. The HIV Neurobehavioral Research Center (HNRC) Group. Journal International Neuropsychological Society 1998;4:291–302.

Miller EN, Bing EG, Selnes OA, Wesch J, Becker JT. The effects of sociodemographic factors on reaction time and speed of information processing. Journal of Clinical and Experimental Neuropsychology 1993;15:66.

O'Bryant SE, Schrimsher GW, O'Jile JR. Discrepancies between self-reported years of education and estimated reading level: Potential implications for neuropsychologists. Applied Neuropsychology 2005;12:5–11. [PubMed: 15788217]

Ostrosky-Solis F, Ardila A, Rosselli M, Lopez-Arango G, Uriel-Mendoza V. Neuropsychological test performance in illiterate subjects. Archives of Clinical Neuropsychology 1998;13:645–660. [PubMed: 14590626]

Roberts RJ, Hamsher KD. Effects of minority status on facial recognition and naming performance. Journal of Clinical Psychology 1984;40:539–545. [PubMed: 6732899]

Ross TP, Lichtenberg PA, Christensen BK. Normative data on the Boston Naming Test for elderly adults in a demographically diverse medical sample. Clinical Neuropsychologist 1995;9:321–325.

Ryan EL, Baird R, Mindt MR, Byrd D, Monzones J, Morgello S. Neuropsychological impairment in racial/ethnic minorities with HIV infection and low literacy levels: Effects of education and reading level in participant characterization. Journal of the International Neuropsychological Society 2005;11:889–898. [PubMed: 16519268]

Stern RA, Singer NG, Silva SG, Rogers HJ, Perkins DO, Hall CD. Neurobehavioral functioning in a nonconfounded group of asymptomatic HIV-seropositive homosexual men. The American Journal of Psychiatry 1992;149:1099–1102. [PubMed: 1636810]

Wechsler, D. Wechsler Adult Intelligence Scale-III. San Antonio, TX: The Psychological Corporation; 1997.

Welsh KA, Fillenbaum G, Wilkinson W, Heyman A, Mohs RC, Stern Y. Neuropsychological test performance in African-American and white patients with Alzheimer's disease. Neurology 1995;45:2207–2211. [PubMed: 8848195]

Wilkinson, GS. Wide Range Achievement Test 3-Administration manual. Delaware: Jastak Associates, Inc.; 1993.

Wilson FL, Racine E, Tekieli V, Williams B. Literacy, readability and cultural barriers: Critical factors to consider when educating older African Americans about anticoagulation therapy. Journal of Clinical Nursing 2003;12:275–282. [PubMed: 12603561]

Table 1

Group characteristics

Variable	Ethnic Group		*p*-value
	Whites Mean (*sd*)	African American Mean (*sd*)	
Age (years)	43 (9.1)	42.27 (7.7)	.65
Self reported education (years)	13.2 (2.6)	13 (2)	.66
WRAT-3 reading grade	11.4 (3)	8.6 (4.4)	<.001
CD4 (median)	157	153	.59
Viral load (median)	6776	10844	.27
Gender (% female)	5.9%	24.2%	.01
Patients identified as NCI (%)	29%	48%	.06

Table 2

Within group comparisons of WRAT-3 and Education-based scores—Whites

Variable	Normative Method		
	Education	**Reading Level**	*p*-value
Trail Making Test–Part A	44.9 (10.7)	45.9 (11)	.02
Trail Making Test–Part B	45.2 (11.4)	46.5 (11.7)	.02
Grooved Pegboard–Dominant	38.8 (12)	39.7 (12)	.01
Grooved Pegboard–Nondominant	36.1 (11.7)	36.6 (11)	.09
Symbol Search	47.5 (10.1)	47.5 (12)	.94
Digit Symbol	46.2 (10.2)	47.7 (11.7)	.08
Letter/Number Sequencing	47.1 (8.9)	47.4 (8.9)	.71
COWAT	47 (11.5)	48.5 (11.6)	.003
PASAT	40.6 (12.4)	44.2 (12.5)	.001

Table 3

Within group comparisons of WRAT-3 and Education-Based Scores—African Americans

Variable	Normative Method		
	Education	**Reading Level**	*p*-value
Trail Making Test–Part A	40.5 (9.7)	42.9 (9.9)	<.001
Trail Making Test–Part B	39.1 (11.1)	42.3 (10.7)	<.001
Grooved Pegboard–Dominant	31.9 (10.5)	33.5 (10.1)	<.001
Grooved Pegboard–Nondominant	32.4 (9.3)	32.7 (9.1)	.155
Symbol Search	46.7 (10.1)	44.1 (9.2)	.001
Digit Symbol	46 (9.8)	47.2 (10.3)	.18
Letter/Number Sequencing	48.4 (8.6)	46.4 (9.3)	.003
COWAT	47.8 (9.2)	45.9 (8.4)	<.001
PASAT	40.6 (7.9)	38.5 (7.3)	<.001

Table 4

Diagnostic accuracy of WRAT-3 and Education-based scores in all participants

	Impairment - WRAT-3	Not Impaired - WRAT-3	Impairment - Education	Non-Impaired - Education
Neurologically symptomatic	46.2% [*]	53.8%	55.8% [*]	44.2%
Neurologically asymptomatic	15.9%	84.1% [†]	22.7%	77.3% [†]
Accuracy of method	65.6%		63.5%	

[*] Sensitivity

[†] Specificity

Table 5

Diagnostic accuracy of WRAT-3 and education-based scores in African Americans

	Impairment - WRAT-3	Not impaired - WRAT-3	Impairment - Education	Non-impaired - Education
Neurologically symptomatic	48.4% [*]	51.6%	61.3% [*]	38.7%
Neurologically asymptomatic	22.2%	77.8% [†]	44.4%	55.6% [†]
Accuracy of method	59%		59%	

[*]
Sensitivity

[†]
Specificity

Table 6

Diagnostic accuracy of WRAT-3 and education-based scores in Whites

	Impairment - WRAT-3	Not impaired - WRAT-3	Impairment - Education	Non-impaired - Education
Neurologically symptomatic	42.9% *	57.1%	47.6% *	52.4%
Neurologically asymptomatic	11.5%	88.5% †	7.7%	92.3% †
Accuracy of method	68.1%		72%	

*
Sensitivity

†
Specificity