Reading level attenuates differences in neuropsychological test performance between African American and White elders

JENNIFER J. MANLY,1,2,3 DIANE M. JACOBS,2,3 PEGAH TOURADJI,3 SCOTT A. SMALL,1,2,3 AND YAAKOV STERN1,2,3,4

1Cognitive Neuroscience Division, Taub Institute for Research on Alzheimer’s Disease and the Aging Brain, Columbia University College of Physicians and Surgeons, New York, NY
2Department of Neurology, Columbia University College of Physicians and Surgeons, New York, NY
3G.H. Sergievsky Center, Columbia University College of Physicians and Surgeons, New York, NY
4Department of Psychiatry, Columbia University College of Physicians and Surgeons, New York, NY

(Received March 6, 2001; Revised August 20, 2001; Accepted August 21, 2001)

Abstract
The current study sought to determine if discrepancies in quality of education could explain differences in cognitive test scores between African American and White elders matched on years of education. A comprehensive neuropsychological battery was administered to a sample of African American and non-Hispanic White participants in an epidemiological study of normal aging and dementia in the Northern Manhattan community. All participants were diagnosed as nondemented by a neurologist, and had no history of Parkinson’s disease, stroke, mental illness, or head injury. The Reading Recognition subtest from the Wide Range Achievement Test–Version 3 was used as an estimate of quality of education. A MANOVA revealed that African American elders obtained significantly lower scores than Whites on measures of word list learning and memory, figure memory, abstract reasoning, fluency, and visuospatial skill even though the groups were matched on years of education. However, after adjusting the scores for WRAT–3 reading score, the overall effect of race was greatly reduced and racial differences on all tests (except category fluency and a drawing measure) became nonsignificant. These findings suggest that years of education is an inadequate measure of the educational experience among multicultural elders, and that adjusting for quality of education may improve the specificity of certain neuropsychological measures. (JINS, 2002, 8, 341–348.)

Keywords: Quality of education, Racial differences

INTRODUCTION
Previous research on cognitive test performance of African Americans has compared African Americans to Whites on intelligence tests (Chen et al., 1994; Heaton et al., 1996; Kaufman et al., 1988; Kush & Watkins, 1997; Overall & Levin, 1978; Reynolds et al., 1987; Vincent, 1991) and screening measures (Bohnstedt et al., 1994; Fillenbaum et al., 1988, 1990; Ford et al., 1996; Kuller et al., 1998; Murden et al., 1991; Teresi et al., 1995; Unverzagt et al., 1996; Welsh et al., 1995). Prior research has also investigated performance of African Americans on neuropsychological tests of naming (Lichtenberg et al., 1994; Roberts & Hamsher, 1984; Ross et al., 1995), reading (Boekamp et al., 1995), 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), and dementia batteries (Carlson et al., 1998; Inouye et al., 1993; Manly et al., 1998a; Marcopulos et al., 1997; Ripich et al., 1997; Unverzagt et al., 1996; Welsh et al., 1995). These racial comparisons have generally shown that despite equivalence on demographic variables such as years of education and socioeconomic status, African Americans obtain lower scores on both verbal and nonverbal cognitive tasks. These discrepancies cause attenuated specificity, such that cognitively normal African Americans are more likely to be mis-
diagnosed as impaired as compared to Whites (Ford-Booker et al., 1993; Klusman et al., 1991; Manly et al., 1998a, 1998b, 1998c; Stern et al., 1992; Welsh et al., 1995).

In each of these studies, investigators used covariance or matching procedures in order to equate racial groups on years of education before comparing neuropsychological test performance, since education often differs in groups of African Americans and Whites. Therefore, these findings depend on whether we can successfully adjust for years of education. However, along with several other authors (Kaufman et al., 1997; Loewenstein et al., 1994; Whitfield & Baker-Thomas, 1999), we argue that matching on quantity of formal education does not necessarily mean that the quality of education received by each racial group is comparable. There is abundant evidence to suggest that years of education is not commensurate between African Americans and Whites and is thus an inadequate estimate of educational experience.

In the United States there is a great deal of discordance between years of education and quality of education; this is especially true among African American elders. Previous studies reveal that African Americans have reading skills significantly below their self-reported education level (Albert & Teresi, 1999; Baker et al., 1996). African Americans educated in the South before the Supreme Court’s 1954 Brown v. Board of Education decision attended segregated schools, which received inferior funding as compared to White Southern schools and most integrated Northern schools (Anderson, 1988). Beginning with the Coleman report (Coleman, 1966), several researchers have demonstrated that school characteristics such as pupil expenditures, teacher quality, pupil/teacher ratios, presence of special facilities such as science laboratories, length of school year/days attended, and peer characteristics were able to account for much of the difference in achievement and other outcomes (e.g., wage earnings), between African Americans and Whites (Hanushek, 1989; Hedges et al., 1994; O’Neill, 1990). The unequal distribution of funds to segregated African American schools in the South in the first half of this century, and the subsequent lower quality of education, was related to lower earnings among African Americans in a number of studies (Margo, 1985; 1990; Smith, 1984; Smith & Welch, 1977; Welch, 1966; 1973). African American children were often employed or used as labor during the harvest, which reduced attendance during the year. This gap in days attended also contributes to differences in quality of education and literacy levels per year in school (Margo, 1985).

In summary, the variable years of education systematically differs between racial groups and is also related to cognitive test performance. If this variable is not commensurate between racial groups, residual confounding will occur and spurious racial differences will be interpreted despite matching groups on years of education. Disparate school experiences, and thus different bases of problem-solving strategies, knowledge, familiarity, and practice could explain why some African Americans obtain lower scores on cognitive measures even after matching groups on years of education. Statistical control of years of education may be inadequate or inappropriate since different scales of measurement are being used between (and within) each racial group (Kaufman et al., 1997; Loewenstein et al., 1994).

Measures of reading achievement correlate with overall academic achievement (Wilkinson, 1993) and with direct measures of quality of education such as pupil expenditures, teacher/student ratios, and teacher education (Hedges et al., 1994). The current study uses the Reading Recognition subtest from the WRAT–3 as an estimate of quality of education, since this measure and other WRAT scores have been found to remain relatively stable over time (Johnstone & Wilhelm, 1996; Wilkinson, 1993).

If the WRAT–3 is used as an estimate of quality of education, it may be important to assure that any correlation between reading score and neuropsychological test performance is independent of variance that might be shared due to the fact that they are both tests. In other words, racial group differences may simply reflect a higher level of “test-wiseness” among non-Hispanic Whites rather than discrepancies in quality of education. Test-wiseness is defined as the ability to use the format and characteristics of a test to achieve a high score (Scruggs & Lifson, 1985), and the use of deduction and item cues to answer questions (Borrello & Thompson, 1985). Test-wiseness has traditionally been measured using multiple-choice items that can only be answered correctly using secondary cues irrelevant to item content.

The purpose of this study was to determine if discrepancies in quality of education could explain differences in cognitive test score between African American and White elders matched on years of education. We hypothesized that we would find significant racial differences on measures of figure memory, verbal abstraction, fluency, and visuospatial skill, but that these differences would be significantly reduced after accounting for our best estimate of quality of education (reading level). We hypothesized that the effect of race would persist after adjusting for an estimate of test-wiseness and that reading level would continue to reduce the effect of race after accounting for test-wiseness.

METHODS

Research Participants

The current sample was selected from participants in the Washington Heights–Inwood Columbia Aging Project (WHICAP), a community-based, epidemiological study of dementia in the ethnically diverse neighborhoods of Northern Manhattan, New York. The WHICAP study follows a random sample of elderly Medicare recipients residing in selected census tracts of Washington Heights and Inwood.

Inclusion/exclusion criteria

All potential participants were age 65 and above and performed the neuropsychological battery in English. Partici-
participants were included if they self-identified their race as White or Black/African American and their “ethnicity” as non-Hispanic according to US Census Criteria (United States Bureau of the Census, 1991). Participants were asked to categorize how well they spoke English: very well; well; or not well. The current study included only those participants who said they spoke English very well. Potential participants were excluded if they had a history of Parkinson’s disease, stroke, head injury with loss of consciousness, alcohol abuse, or serious mental illness such as depression or schizophrenia. Only WHICAP participants who showed no neurological or functional signs of dementia were included in the current analyses. This determination was made on the basis of a physician’s clinical examination, which included a rating of daily functioning (see procedures below). The physician’s diagnosis was used as a gold standard for the absence of dementia, since the neurological assessment was made independent of the participant’s performance on the neuropsychological battery.

Medical evaluation

A physician recorded medical history and medications in a semi-structured format. Neurological and brief physical examinations were performed, including assessment of extrapyramidal signs. Functional status was measured using Part 1 of the Blessed Dementia Rating Scale (BDRS; Blessed et al., 1968) and the Schwab and England rating scale of activities of daily living (Boller et al., 1980). From this information, the physician determined whether the participant met criteria for delirium or dementia using Diagnostic and Statistical Manual of Mental Disorders–Revised Third Edition (American Psychiatric Association, 1987) criteria.

Procedure

Neuropsychological battery

The neuropsychological measures used in the current study were selected to assess cognitive functions that are typically affected in dementia and have been shown to effectively distinguish between normal aging and dementia in this community (Stern et al., 1992). The evaluation included measures of learning and memory, orientation, abstract reasoning, language, and visuospatial ability. Specific ability areas and tests administered include verbal list learning and memory [Selective Reminding Test (SRT); Buschke & Fuld, 1974], nonverbal memory [multiple choice version of the Benton Visual Retention Test (BVRT); Benton, 1955], orientation [items from the Mini Mental State Examination (MMSE); Folstein et al., 1975], verbal reasoning [Similarities subtest of the Wechsler Adult Intelligence Scale–Revised (WAIS–R); Wechsler, 1981], nonverbal reasoning (Identities and Oddities subtest of the Mattis Dementia Rating Scale; Mattis, 1976), naming (15-item version of the Boston Naming Test; Kaplan et al., 1983), letter fluency (Controlled Word Association; Benton & Hamsher, 1976), category fluency [animals, food, and clothing, using procedures from the Boston Diagnostic Aphasia Examination (BDAE); Goodglass & Kaplan, 1983], repetition (high-frequency phrases of the BDAE; Goodglass & Kaplan, 1983), auditory comprehension (first six items of the Complex Ideational Material subtest of the BDAE; Goodglass & Kaplan, 1983), visuconstruction (Rosen Drawing Test; Rosen, 1981), and visuoperceptual skills (multiple choice matching of figures from the BVRT; Benton, 1955).

Reading level

Reading level was measured using the Reading Recognition subtest from the Wide Range Achievement Test–Version 3 (Wilkinson, 1993). Participants were asked to name letters and pronounce words out of context. The words are listed in order of decreasing familiarity and increasing phonological complexity. Consistent with the standard instructions for administration, a basal of 5 correct and a ceiling of 10 incorrect was used. WRAT–3 grade equivalent scores were derived from the normative values for people age 70 to 75 years from the manual.

Test-wiseness

Word Recognition score from the Selective Reminding Test (SRT) was used as an estimate of test-wiseness. SRT Word Recognition requires the participant to select previously presented words from four-choice multiple-choice arrays. Distractors are both phonemic and semantic in nature (e.g., correct choice = helmet; distractors = armor, velvet, bacon).

Statistical Methods

In order to create groups matched on years of education and sex, a stratified random sampling method was applied. The sample was stratified by race (White, African American), sex (male, female) and four categories based on years of education (0–8; 9–11; 12–15; and greater than 16 years). Within each Education × Race × Sex cell, a random sample of subjects was selected using the SPSS (1998) sample function. Equal numbers of participants were selected within each Education × Race × Sex stratum.

Chi-square tests were used to compare African Americans and Whites on estimated grade level based on their WRAT–3 reading performance, and t tests were used to compare raw WRAT scores by race.

A MANOVA was used to compare African Americans and Whites on the overall test battery. Thirteen measures from the test battery were included as dependent variables: total raw scores for immediate recall and delayed recall from the SRT; BVRT matching and recognition memory; items correct on the MMSE Orientation; 15-item Boston Naming; BDAE repetition; BDAE comprehension; Mattis Identities and Oddities; number correct on the Rosen Drawing Test; raw score on WAIS–R Similarities; and mean number of words generated over three 60-s trials for category and letter fluency. Post-hoc ANOVAs were performed to
determine which specific neuropsychological measures contributed significantly to the overall multivariate effect. In order to strike a balance between the likelihood of committing Type I and Type II errors, an alpha level of $p < .01$ was used to determine statistically significant univariate differences.

WRAT–3 reading score was then added as an additional covariate into a MANCOVA, in order to test the effect of race after adjusting for our estimate of quality of education. Again, an alpha level of $p < .01$ was used to determine statistically significant univariate differences. A final MANCOVA was performed in order to determine if covarying for SRT score had the same effect on cognitive test differences between African American and Whites after accounting for SRT recognition performance, our estimate of test-wiseness.

**RESULTS**

**Sample Characteristics**

A total of 708 English-speaking WHICAP participants had complete neuropsychological evaluations, WRAT–3 data, and neurological exams. Of these, 322 identified as non-Hispanic White and 386 identified as non-Hispanic Black. After excluding participants who did not meet inclusion/exclusion criteria, 248 White and 317 African American participants remained.

Whites and African Americans were compared on years of education and age using $t$-tests, and chi-square analysis compared the proportion of women in each racial group. As expected, there was a significant difference in years of education between African Americans ($M = 11.0$, $SD = 3.6$) and Whites ($M = 13.8$, $SD = 3.2$; $t(563) = 9.4$, $p < .001$). There was no age difference between African American ($M = 74.5$, $SD = 5.9$) and White elders ($M = 75.0$, $SD = 5.9$; $t(563) = 0.05$, $p = .96$). The White group was 60% female, and 76% of the African American elders were women [$\chi^2(1, N = 565) = 18.0$, $p < .001$]. A stratified random sampling method (described above) was then used to match the White and African American participants on years of education and sex.

Groups of 192 African American and 192 White elders matched on educational attainment and sex distribution were formed through the stratified random sampling procedure. Table 1 compares the demographic characteristics of the two groups. As directed by the matching procedure, both groups were 68% female, and there was no significant difference in years of education between African Americans and Whites. There was still no age difference between the groups.

**Race, Reading Level, and Neuropsychological Test Performance**

As expected, there was a significant difference in WRAT–3 reading scores between racial groups (Table 1). When raw WRAT–3 scores were converted into grade-equivalent scores and compared with reported years of education, self-reported years of education was an overestimate of actual reading level for a higher proportion of African Americans (33%) than Whites (7%).

A MANCOVA revealed that African Americans scored significantly lower on the neuropsychological test battery overall as compared to Whites [$F(13,370) = 6.2$, $p < .001$]. Follow-up univariate tests revealed significant differences between African Americans and Whites on measures of word list learning and memory (SRT total immediate and delayed recall), figure memory (BVRT recognition), abstract reasoning (WAIS–R Similarities and DRS identities and oddities), letter fluency, category fluency, and visuospatial skill (Rosen drawings and BVRT matching). Table 2 shows racial group means and standard deviations on each test within the battery.

WRAT–3 scores accounted for 34% of the variance of test scores overall, which was significant [$F(13,369) = 14.3$, $p < .001$]. When WRAT–3 scores were entered into the multivariate model, the overall effect of race remained significant [$F(13,368) = 2.7$, $p = .001$], but the effect size was significantly reduced. Follow-up univariate comparisons

<table>
<thead>
<tr>
<th>Variable</th>
<th>African American $M$ (SD)</th>
<th>White $M$ (SD)</th>
<th>Statistic $t$ or $\chi^2$</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>N</td>
<td>192</td>
<td>192</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Female</td>
<td>68.2%</td>
<td>68.2%</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td>Age</td>
<td>73.9 (5.8)</td>
<td>74.6 (5.9)</td>
<td>1.0</td>
<td>.30</td>
</tr>
<tr>
<td>Years of education</td>
<td>12.8 (2.8)</td>
<td>13.0 (3.0)</td>
<td>.60</td>
<td>.55</td>
</tr>
<tr>
<td>WRAT–3 reading score</td>
<td>44.2 (7.2)</td>
<td>49.3 (4.1)</td>
<td>8.4</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Reading level = reported grade</td>
<td>47%</td>
<td>56%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading level &gt; reported grade</td>
<td>20%</td>
<td>38%</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading level &lt; reported grade</td>
<td>33%</td>
<td>7%</td>
<td>48.98</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>
Table 2. Effect of race and WRAT-3 reading score on neuropsychological test score

<table>
<thead>
<tr>
<th>Test</th>
<th>African American</th>
<th>White</th>
<th>Effect of race</th>
<th>After covarying for WRAT reading score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>p</td>
<td>F</td>
<td>p</td>
</tr>
<tr>
<td>Learning/Memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SRT total recall</td>
<td>39.8 (10.1)</td>
<td>43.5 (10.1)</td>
<td>12.9</td>
<td>.000</td>
</tr>
<tr>
<td>SRT delayed recall</td>
<td>5.8 (2.7)</td>
<td>6.7 (2.9)</td>
<td>8.9</td>
<td>.003</td>
</tr>
<tr>
<td>BVRT recognition memory</td>
<td>7.4 (1.8)</td>
<td>8.1 (1.5)</td>
<td>20.4</td>
<td>.000</td>
</tr>
<tr>
<td>Orientation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MMSE orientation</td>
<td>9.7 (0.7)</td>
<td>9.8 (0.5)</td>
<td>3.5</td>
<td>.062</td>
</tr>
<tr>
<td>Abstract reasoning</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WAIS-R Similarities raw</td>
<td>12.0 (6.9)</td>
<td>16.0 (6.3)</td>
<td>34.8</td>
<td>.000</td>
</tr>
<tr>
<td>DRS Identities &amp; Oddities</td>
<td>14.6 (1.7)</td>
<td>15.1 (1.3)</td>
<td>10.8</td>
<td>.001</td>
</tr>
<tr>
<td>Language</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boston Naming</td>
<td>14.0 (1.4)</td>
<td>14.1 (1.5)</td>
<td>0.5</td>
<td>.480</td>
</tr>
<tr>
<td>Letter fluency</td>
<td>9.9 (3.8)</td>
<td>12.2 (4.1)</td>
<td>31.8</td>
<td>.000</td>
</tr>
<tr>
<td>Category fluency</td>
<td>14.6 (3.8)</td>
<td>16.8 (3.8)</td>
<td>31.4</td>
<td>.000</td>
</tr>
<tr>
<td>BDAE repetition</td>
<td>7.8 (0.6)</td>
<td>7.8 (0.5)</td>
<td>0.0</td>
<td>.848</td>
</tr>
<tr>
<td>BDAE comprehension</td>
<td>5.5 (0.9)</td>
<td>5.8 (0.7)</td>
<td>13.4</td>
<td>.000</td>
</tr>
<tr>
<td>Visuospatial ability</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rosen drawing</td>
<td>2.6 (0.9)</td>
<td>3.0 (0.8)</td>
<td>30.0</td>
<td>.000</td>
</tr>
<tr>
<td>BVRT matching</td>
<td>8.9 (1.4)</td>
<td>9.4 (1.2)</td>
<td>13.1</td>
<td>.000</td>
</tr>
</tbody>
</table>

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

Table 2 showed that after accounting for WRAT–3 score, racial differences on all measures except category fluency and the Rosen drawing test were no longer significant.

Next, our best estimate of “test-taking savvy,” the SRT Word Recognition task, was added as a predictor of neuropsychological test scores in a multivariate model. SRT Recognition was related to scores on all ten measures except the Rosen drawing test (p < .001 for all); however, even after SRT Recognition was accounted for, there was no change in the overall effect of race on cognitive test score [F(13,369) = 5.7, p < .001]. When WRAT–3 score was added as a predictor in the model in addition to SRT Recognition score, the effect of race became nonsignificant for all measures except Rosen drawing and category fluency, the same result as when WRAT–3 alone was used as a covariate.

DISCUSSION

The current study found that accounting for WRAT–3 Reading recognition performance, an estimate of quality of education, attenuates racial group differences on most cognitive tests between African American and White elders matched on years of education. These findings suggest that the full extent of discrepancies in educational experience between African Americans and Whites are not captured by a simple “highest-grade-attained” variable, and thus residual confounding may explain findings of persistent race effects after matching groups on years of education.

Although adjustment for reading recognition accounted for racial differences on the letter fluency task, the effect of race on category fluency remained significant. In addition, WRAT–3 score did not fully account for racial differences in the Rosen drawing task, but discrepancies in figure matching became nonsignificant. In other words, WRAT–3 score was significantly correlated with Rosen drawing and category fluency in both racial groups and accounted for racial differences in related tasks; however, differences in reading level cannot fully explain racial discrepancies on these two measures. It is possible that other cultural differences, unmeasured in this study, can account for racial differences in performance on these two measures. One possible factor is cultural experience, which has been shown to account for some test differences between African American and White HIV-positive individuals (Manly et al., 1998d). It is also possible that WRAT–3 score does not capture the aspects of quality of education that are related to semantic fluency and copying geometric figures. Future study will determine whether variables such as per student costs, teacher salaries, or length of school year can more completely explain discrepancies in test performance between these two racial groups.

Although the utility of single-word reading measures such as the WRAT in premorbid IQ estimation is controversial (Dura et al., 1989; Johnstone et al., 1996; Johnstone & Wilhelm, 1996; O’Carroll, 1995; Storandt et al., 1995), one longitudinal study of incident dementia (Schmand et al., 1998) concluded that reading level remains a valid estimate
of premorbid ability in mild and questionable dementia; this is the same context in which the results of this study will be most useful. Certainly, a major limitation of reading recognition scores as an estimate of educational quality is that scores would be inaccurate among elders with history of learning disabilities or severe language disturbance related to a neurological disease.

Educational quality is only one of many variables that may influence reading recognition scores, including general cognitive ability, reading experience obtained through access to books in the home or as a result of occupational demands, and test-wiseness. We attempted to control for one of these factors by determining if reading score attenuated the effect of race after accounting for an estimate of test-wiseness. SRT Word Recognition score was used as an estimate of test-taking ability because it is a multiple-choice task in which participants may use a process of elimination and regularities of distractors to help narrow the field of possible correct answers. This score is certainly not a perfect measure of test-wiseness; rather, its inclusion served to adjust the neuropsychological test scores for shared attributes of all cognitive tasks.

The results of the current study have both practical and conceptual implications. Practically, these findings suggest that race-specific norms which correct for years of education may be less accurate than norms which correct for quality of education and are not specific to one racial classification. Neuropsychologists use demographically corrected norms to improve the accuracy of predictions of a normal individual’s performance, and thus improve the specificity and sensitivity of a task to detect cognitive impairment. Just as age and sex are expected to adjust expectations of an individual’s performance, years of education has traditionally been used to adjust for changes in baseline knowledge, strategy, and skill that are accompanied by formal schooling. However, this study adds to the multitude of evidence, collected mainly by educational psychologists and economists, that shows racial inequity in the value of any unit of quantity of education. Based on the results of this study, we propose that regardless of race, the WRAT–3 reading test measures educational experience more accurately than years of education, and thus is a superior assessment of the knowledge, strategy, and skills needed to perform well on traditional neuropsychological tasks. Test scores adjusted for reading level can be used to predict performance more accurately than if only years of education and racial classification were used.

The current research approach also addresses the fallacy underlying most comparisons of test performance between African Americans and Whites; the assumption that because racial groups are matched on socioeconomic variables such as years of education, persistent race effects can be interpreted as biologically meaningful. One way to avoid this faulty conclusion is to specify and measure the experimental, attitudinal, or behavioral variables that are expected to distinguish those belonging to different ethnic groups, and which also vary among individuals within an ethnic group (Helms, 1992). These factors may reflect underlying reasons for the relationship between racial classification and cognitive test performance. Because racial differences on cognitive test performance were significantly reduced (and in most cases, eliminated) by accounting for just one indicator of educational quality, the current findings support this approach. This investigational approach may illuminate factors which can explain not only ethnic group differences on cognitive tests, but can inform our future development of measures designed to measure cognitive abilities salient within African American culture. The effect of specific cultural and educational factors on cognitive test performance must be well understood before we attempt to develop culture-fair measures.

ACKNOWLEDGMENTS

This research was supported by federal grants AG16206 (J. Manly), AG07232 (R. Mayeux), the Alzheimer’s Association, and the New York City Speakers Fund for Biomedical Research—Toward the Science of Patient Care. The authors thank Rosann Costa for her help with data management and Maria Gonzalez-Diaz, Cherita McDowell, and Jules Fleurimont for their assistance with scheduling and interviewing participants.

REFERENCES


veterans using the American version of the National Adult Reading Test. *Journal of Clinical and Experimental Neuropsychology,* 17, 645–653.


