Journal of the International Neuropsychological Society (2012), **18**, 277–285.

doi:10.1017/S1355617711001688

Quality, and not Just Quantity, of Education Accounts for Differences in Psychometric Performance between African Americans and White Non-Hispanics with Alzheimer's Disease

Alexander L. Chin,[1] Selam Negash,[2] Sharon Xie,[3] Steven E. Arnold,[2] AND Roy Hamilton[2,4]

[1]Perelman School of Medicine, University of Pennsylvania, Philadelphia, Pennsylvania
[2]Penn Memory Center, University of Pennsylvania, Philadelphia, Pennsylvania
[3]Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, Pennsylvania
[4]Department of Neurology, University of Pennsylvania, Philadelphia, Pennsylvania

(RECEIVED January 28, 2011; FINAL REVISION November 7, 2011; ACCEPTED November 7, 2011)

Abstract

The effect of race on cognitive test performance in the evaluation of Alzheimer's disease (AD) remains controversial. One factor that may contribute substantially to differences in cognitive test performance in diverse populations is education. The current study examined the extent to which quality of education, even after controlling for formal years of education, accounts for differences in cognitive performance between African Americans and White Non-Hispanics (WNHs). The retrospective cohort included 244 patients diagnosed with AD who self-identified as African Americans ($n = 51$) or WNHs ($n = 193$). The Wechsler Test of Adult Reading (WTAR) was used as an estimate of quality of education. In an analysis that controlled for traditional demographics, including age, sex, and years of formal education, African Americans scored significantly lower than WNHs on the Mini-Mental State Examination, as well as on neuropsychological tests of memory, attention, and language. However, after also adjusting for reading level, all previously observed differences were significantly attenuated. The attenuating effect remained even after controlling for disease severity, indicating that reading scores are not confounded by severity of dementia. These findings suggest that quality, and not just quantity, of education needs to be taken into account when assessing cognitive performance in African Americans with AD. (*JINS*, 2012, *18*, 277–285)

Keywords: Cross-cultural comparisons, Neuropsychological tests, Aging/psychology, Dementia, European Continental Ancestry Group/psychology, African Americans/psychology, Reading, Diagnostic errors/psychology, Bias (epidemiology), Socioeconomic factors, Philadelphia

INTRODUCTION

Current diagnosis of Alzheimer's disease (AD) relies largely upon the clinical judgment of physicians. The lack of widely available, validated imaging or biochemical biomarkers for the detection of AD pathology, particularly in underserved populations, necessitates the use of neuropsychological instruments as the foundation of dementia screening and diagnosis. Traditional measures of demographics, including age, sex, and number of years of education, are typically used to control for results obtained from these tests. However, research has suggested that even after controlling for these variables, African Americans score significantly lower on a variety of measures compared to their White Non-Hispanic (WNH) counterparts (Fillenbaum, Huber, & Taussig, 1997; Welsh et al., 1995; Wood, Giuliano, Bignell, & Pritham, 2006).

If ethnoracial differences exist in performance on neuropsychological tests used to screen for dementia, cognitively normal African Americans may be more likely to receive a misdiagnosis of AD than age- and education-matched WNHs (Manly et al., 1998; Pedraza & Mungas, 2008). For example, researchers have found that the Mini-Mental State Examination (MMSE) may have up to a 42% false-positive rate for cognitive impairment among African Americans as

Correspondence and reprint requests to: Roy Hamilton, Penn Memory Center, 3615 Chestnut Street, Philadelphia, PA 19104. E-mail: roy.hamilton@uphs.upenn.edu

compared to a 6% rate among WNHs (Stephenson, 2001). Similar between-group score discrepancies have been found on measures of verbal and nonverbal learning and memory, abstract reasoning, language, and visuospatial skill (Byrd, Touradji, Tang, & Manly, 2004; Manly et al., 1998; Manly, Touradji, Tang, & Stern, 2003).

The issue of misdiagnosis of AD within diverse populations is complicated by evidence that the rates of dementia among African Americans and Latinos may be higher than among WNHs (Demirovic et al., 2003; Tang et al., 2001). Whether or not these differences can be attributed to true differences in the pathophysiology of AD between ethnoracial groups, to potentially elevated false-positive rates of screening tools in diverse populations, or to some other yet unidentified factor remains to be determined. To elucidate the differential role of pathophysiology in the development of AD within minority populations, neuropsychological instruments used to screen for dementia must be properly controlled for environmental factors such as education, as well as sociocultural biases of testing content and process. Because formal education strongly influences cognitive test performance (Tombaugh & McIntyre, 1992), educational experience should be one of the primary factors controlled for in the assessment of cognition.

Previous research has suggested that quality of education may be a more accurate indicator of educational experience than the number of years of education (Manly, Jacobs, Touradji, Small, & Stern, 2002). Even though African Americans and WNHs may be matched on years of formal education, the quality of education across ethnoracial groups may still differ due to socioeconomic disparities, historically segregated educational systems, and other factors (Whitfield & Wiggins, 2003). As a result, neuropsychological test scores may not be properly adjusted for the true effect of education if performance is controlled for solely based on years of schooling.

Research by Manly and colleagues (1998, 2002, 2003) in this area has focused on cognitively normal, non-demented elderly populations. However, to our knowledge, no studies performed to date have examined the effect of quality of education on neuropsychological test performance in a cognitively impaired population. As a form of retrospective analysis, this research is important to help elucidate whether or not African Americans have already been disproportionately diagnosed with AD on the basis of inadequately controlled and inappropriately poor test performance. In addition, since neuropsychological tests are used to diagnose AD in cognitively impaired individuals, it should be established that quality of education is an appropriate control measure in both cognitively normal and cognitively impaired populations.

In this retrospective cohort study, we examined the impact of both years of education and quality of education on ethnoracial differences in cognitive test performance between African Americans and WNHs with AD. We hypothesized that quality of education would better account for differences between African American and WNH patients than years of education.

METHODS

Research Participants

The University of Pennsylvania Alzheimer's Disease Center (ADC) works alongside a University-associated primary care practice, the Penn Memory Center (PMC). Located in West Philadelphia, the ADC evaluates a diverse cohort of patients from local communities, including primarily low-income, African American and Latino communities. In accord with University-approved IRB protocols and with the informed consent of patients, the ADC acquires data that includes demographic information, quality of life measures, neuropsychological test performance, and biomarker data from neuroimaging, cerebrospinal fluid (e.g., CSF tau and amyloid beta levels) and blood (apolipoprotein E genotype status). Using standardized clinical criteria and laboratory data, experienced clinicians, including neurologists, psychiatrists, and neuropsychologists establish a consensus diagnosis for each patient.

Patients were included in the analysis if they: (1) were diagnosed with probable AD by consensus diagnosis, (2) were evaluated with the full psychometric battery detailed below, (3) had an age greater than 50 at time of initial evaluation, and (4) had a self-reported ethnoracial classification of WNH or non-Hispanic Black/African American (referred to as "African American"). Of 3223 patients evaluated at the PMC between December 1989 and April 2010, 244 patients met these inclusion criteria.

Clinical Assessments

All clinical data used in this study were collected at each participant's initial visit to the PMC.

Demographics and clinical characteristics

Participants or their reliable informants completed a questionnaire for age, sex, self-reported ethnoracial group membership, years of education, age at onset of cognitive impairment, interval between symptom onset and first visit to the PMC, current living situation, marital status, and handedness. Past medical history was documented for diabetes mellitus, hypertension, hyperlipidemia, and/or cardiovascular, cerebrovascular or peripheral vascular disease.

The Wechsler Test of Adult Reading (WTAR)

This is a 50-item assessment that requires the reading of irregularly spelled words to assess recognition and prior learning of the word (Wechsler, 2001). It serves as a proxy for educational quality, and correlates highly with tests of reading and achievement (Wechsler, 2001). Normative standard scores are adjusted for demographics and are co-normed with the WAIS-III (Wechsler, 1997a) and the WMS-III (Wechsler, 1997b) to predict performance on these assessments. The test also displays good internal consistency and test–retest reliability (Wechsler, 2001).

https://doi.org/10.1017/S1355617711001688 Published online by Cambridge University Press

Cognition and dementia severity

Cognition was assessed using a neuropsychological battery merging components of the Consortium to Establish a Registry for Alzheimer's Disease (CERAD) (Morris et al., 1989), Alzheimer's Disease Centers' Uniform Data Set (UDS) (Weintraub et al., 2009) and PMC test protocols. This battery included the MMSE (Folstein, Folstein, & McHugh, 2004), Logical Memory (Story A) and Digit Span subtests from the Wechsler Memory Scale-Revised (WMS-R) (Wechsler, 1987a), a 10-item word list memory task (including word list recall and word list recognition) (Morris et al., 1989), category (Morris et al., 1989) and letter fluency tasks, 30-Item Boston Naming Test (Kaplan, Goodglass, & Weintraub, 1983), the 30-odd numbered items, Trail Making Test (Reitan & Wolfson, 1993), Digit Symbol subtest from the Wechsler Adult Intelligence Scale-Revised (WAIS-R) (Wechsler, 1987b), Constructional Praxis and Praxis Recall (Morris et al., 1989), clock draw and clock copy. A CERAD total score (maximum score = 100) was attained using the methods previously described by Chandler and colleagues (2005), and included the following tests: Verbal Fluency, Boston Naming Test, Word List Learning, Constructional Praxis, Word List Recall, and Word List Recognition.

The Dementia Severity Rating Scale (DSRS) (Clark & Ewbank, 1996; Xie et al., 2009) and the Clinical Dementia Rating scale (CDR) (Morris, 1997) provided global measures of dementia severity, while neuropsychiatric symptoms were evaluated using the Neuropsychiatric Inventory Questionnaire (NPI-Q) (Kaufer et al., 2000).

Functional severity

Functional impairment was examined using the Pfeffer Functional Assessment Questionnaire (FAQ) (Pfeffer, Kurosaki, Harrah, Chance, & Filos, 1982). The FAQ is a 10-item informant-report questionnaire that inquires into an older adult's ability to manage finances, complete forms, shop, perform games of skill or hobbies, prepare hot beverages, prepare a balanced meal, follow current events, attend to television programs, books or magazines, remember appointments, and travel out of the neighborhood. Impairment on each item is graded 0–3, a higher score suggesting a more impaired state of functional activity. The FAQ has good reliability (item-total correlations $\geq .80$) and validity (correlations $\geq .70$ with measures of mental status, daily function, and clinical diagnosis) (Pfeffer et al., 1982).

Mood

Mood symptoms were quantified using the Geriatric Depression Scale (GDS) (Yesavage et al., 1982).

Apolipoprotein E (APOE)

APOE genotyping was performed using DNA collected via phlebotomy and extracted from peripheral leukocytes (Petersen et al., 1995).

Statistical Methods

To evaluate differences in the demographic and cognitive variables between WNHs and African Americans, *t* tests were used to compare continuous variables (e.g., age at initial visit, number of years of education, and WTAR standard score), while χ^2 tests were used to compare categorical variables (e.g., sex and family history of dementia).

Hierarchical linear regression modeling was conducted, in which covariates were entered into blocks. In the first block, age at initial visit and sex were entered. Second block added years of education. The third block added WTAR standard score (in addition to age, sex, and years of education) to examine the change in the effect of ethnoracial classification before and after controlling for WTAR score. The correlation between WTAR score and race was examined through a regression analysis to rule out the possibility of collinearity between race and WTAR. To adjust for multiple testing in this study, statistical significance was determined by an alpha level of $p < .01$. All statistical tests were two-sided. Statistical analyses were performed using IBM PASW Statistics 18P (2009).

RESULTS

Demographic Characteristics

A total of 244 patients with a consensus diagnosis of probable AD were included in the analysis. Of these, 51 self-identified as African American and 193 as WNH. Females were over-represented in both ethnoracial groups but the proportion of WNH females (66.3%) did not differ significantly from the proportion of African American females (72.5%). The two groups differed significantly with respect to years of education and WTAR standard score. The distribution of WTAR score by race is shown in Figure 1.

Compared to WNHs, African Americans had fewer years of education (12.2 years *vs.* 14.2) and lower WTAR score (84.2 *vs.* 105.0). The correlation of WTAR score with years of education was significant for both African Americans, $r = .67$, $p < .0001$ and WNHs, $r = .70$, $p < .0001$. There was no significant difference in age at initial evaluation between groups. These results are summarized in Table 1.

Cognitive and Behavioral Characteristics

African Americans diagnosed with AD presented to the Penn ADC with more severe cognitive and neuropsychiatric impairment than WNHs as measured by the Dementia Severity Rating Scale (DSRS), the Clinical Dementia Rating (CDR) Scale, and the Neuropsychiatric Inventory Questionnaire (NPI-Q). However, no significant differences were found between groups with regard to age of onset of cognitive symptoms, Geriatric Depression Scale (GDS) score, Functional Assessment Questionnaire (FAQ) score, or APOE ε4 genotype. Although some prior evidence has suggested that minority patients often receive delayed care for dementia

https://doi.org/10.1017/S1355617711001688 Published online by Cambridge University Press

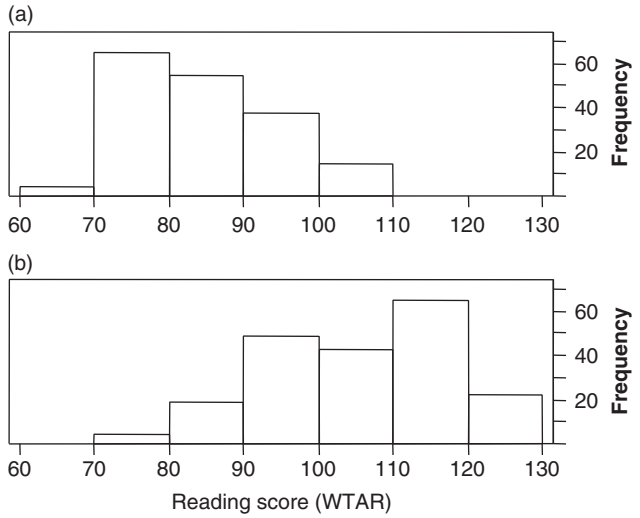

Fig. 1. Distribution of WTAR score by race: (a) African Americans, (b) White Non-Hispanics (WNH).

(Manly, Jacobs, & Mayeux, 1999), the average interval between reported onset of cognitive symptoms and initial evaluation at the ADC did not differ significantly between African Americans and WNHs. These results are summarized in Table 2.

Regression Analysis

To rule out the possibility of collinearity between race and WTAR score, we ran a regression model using WTAR score as the response and race as the independent variable. The R-squared for the model was 0.355, suggesting that collinearity does not exist between race and WTAR score. In addition, linear regression modeling showed that WTAR score itself was significantly correlated with both race ($p < .001$) and scores of cognition (as measured by the CERAD composite score; $p < .001$), indicating that WTAR is a potential confounding variable when examining the effect of race on cognitive measures.

After controlling for age at initial visit and sex, significant differences existed between African Americans and WNHs on the MMSE, word list recall, word list recognition, Boston Naming Test, constructional praxis, constructional praxis recall, digit symbol, reverse Digit Span, Trail Making Test-Part A, and CERAD composite score. Linear regression modeling revealed that African Americans scored significantly lower on each of these tests than age- and sex-matched WNHs ($p < .01$). After adding years of education as an additional covariate to each linear regression model, all test score differences between African Americans and WNHs that were previously found to be significant were attenuated. However, significant between-group performance differences still existed on the MMSE, CERAD composite, word list recognition, Boston Naming Test, constructional praxis recall, backwards digit span, and Trail Making Test A.

As a final step in our linear regression modeling, WTAR standard score was added as a fourth control factor to age, sex, and years of education. The results are summarized in Table 3.

After controlling for WTAR score, no significant performance difference existed between African Americans and WNHs on any neuropsychological test included in the analysis. WTAR score fully attenuated the effect of ethnoracial classification on neuropsychological test performance. Moreover, after controlling for WTAR score, number of years of education was no longer a significant predictor of many of the test scores (data not shown), suggesting that WTAR score alone may explain much of the variation in neuropsychological test performance associated with education. Table 4 shows the adjusted means for the test scores before and after adjusting for WTAR, and demonstrates the significant attenuations of ethnoracial differences after reading score had been entered in the model.

As also suggested by our reviewers, an alternative explanation to the above findings, however, is that WTAR effects are confounded by differences in functional impairment. That is, since African Americans in our study appear to have more severe impairment compared to WNHs, it is likely that the observed effects are due to differences in functional severity rather than to reading ability. To address this concern, we performed two additional subanalyses accounting for functional status and disease duration. In the first, we matched the two ethnoracial groups on functional severity using the Functional Assessment Questionnaire (FAQ). This analysis yielded 190 subjects, 51 AAs and 139 WNHs (AA: mean = 16.1; $SD = 8.9$; WNH: mean = 16.1; $SD = 5.4$). We then repeated the above regression analyses in these subgroups that were matched on functional severity. Table 5 summarizes the results. As the table indicates, the WTAR score remained to have attenuating effect

Table 1. Demographic characteristics

Variable	African American *M* (*SD*)	WNH *M* (*SD*)	Statistic *t* or χ^2	*p*
N	51	193		
Sex (% female)	72.5%	66.3%	0.715	.398
Family history of dementia (% Yes)	29.2%	38.7%	1.495	.221
Age at initial evaluation	76.9 (6.5)	74.6 (8.4)	−1.821	.070
Years of education	12.2 (3.1)	14.2 (3.1)	4.141	<.001*
WTAR standard score	84.2 (9.0)	105.0 (12.0)	11.516[a]	<.001*

Note. WTAR = Wechsler Test of Adult Reading.
[a]Equal variances not assumed, measured using Levene's Test for Equality of Variances, $p < .01$.
*Statistically significant to alpha level of .01.

https://doi.org/10.1017/S1355617711001688 Published online by Cambridge University Press

Table 2. Cognitive, psychiatric, and genetic characteristics

Variable	African American *M* (*SD*)	WNH *M* (*SD*)	Statistic *t* or χ^2	*p* value
Age of onset	74.0 (6.7)	71.9 (8.7)	−1.899	.061
Interval between onset and evaluation (yr)	3.1 (3.6)	2.7 (2.1)	−1.111	.268
DSRS	16.8 (9.3)	12.8 (6.5)	−2.877[a]	.005*
CDR sum of boxes	6.9 (3.5)	4.5 (2.3)	−4.413[a]	<.001*
GDS	3.2 (3.2)	2.4 (2.5)	−1.755	.081
FAQ	16.1 (9.0)	12.6 (7.1)	−2.481[a]	.016
NPI-Q	4.0 (2.8)	2.9 (2.2)	−2.965	.003*
APOE ε4%	62.5%	57.0%	0.382	.536

Note. DSRS = Dementia Severity Rating Scale; CDR = Clinical Dementia Rating; GDS = Geriatric Depression Scale; FAQ = Functional Assessment Questionnaire; NPI-Q = Neuropsychiatric Inventory Questionnaire; APOE = Apolipoprotein E.
[a]Equal variances not assumed, measured using Levene's Test for Equality of Variances, $p < .01$.
*Statistically significant to alpha level of .01.

even in the sample that was matched on functional severity. While adjusting for years of education significantly attenuated ethnoracial differences in most of the test scores, there were still significant group differences in several tests that were only attenuated after adjusting for WTAR. These were: MMSE, CERAD composite, word list recognition, and Boston Naming Test. These findings suggest that the attenuating effect of reading scores is not confounded by functional severity in that adjusting for WTAR score in the severity-matched sample still led to significant attenuations of initially observed ethnoracial differences in psychological performance.

In the second subanalysis, we stratified our sample by disease duration to further ascertain whether the observed WTAR effects are confounded by severity of the disease. The patients were grouped into short (<4 years) and long (≥4 years) duration. This yielded 240 subjects with disease duration data, 175 of whom had short duration (34 AA and 141 WNH) and 65 had long duration (15 AA and 50 WNH). Even though we performed regression analyses for both durations, we only report data from the short duration since the small sample size in the long duration subgroup limits valid interpretation of results. In this analysis, tests that showed significant ethnoracial differences after adjusting for education were: MMSE, CERAD composite, word list recognition, and Boston Naming Test, Praxis Recall, and Trails A. Notably, these were the same tests that remained significant after adjusting for education in our original sample. In the subsequent analysis that adjusted for WTAR score,

Table 3. Linear regression models depicting ethnoracial differences in neuropsychological performance

	N	AA-WNH coeff (w/o education or WTAR)[a]	*p* value	AA-WNH coeff (w/ education, w/o WTAR)[a]	*p* value	AA-WNH coeff (w/ education and WTAR)[a]	*p* value
MMSE	244	−3.579	<.001*	−2.710	<.001*	−0.228	.783
CERAD total	239	−9.822	<.001*	−7.819	.0003*	−0.547	.829
WL memory	239	−0.412	.162	−0.246	.412	0.210	.573
WL delayed recall	239	−0.679	.008*	−0.589	.026	−0.068	.834
WL recognition	238	−2.546	<.001*	−2.254	<.001*	−1.371	.026
LM immediate	235	−0.345	.501	0.075	.884	0.992	.118
LM delay	236	−0.502	.071	−0.343	.225	0.187	.588
Boston naming	242	−6.416	<.001*	−5.481	<.001*	−1.392	.266
Category Fluency	242	−1.315	.057	−0.942	.180	0.483	.576
Clock draw	235	0.968	.051	0.714	.160	−0.251	.689
Clock copy	236	0.687	.072	0.341	.372	−0.856	.063
Praxis construction	241	−1.077	.001*	−0.766	.023	−0.009	.501
Praxis recall	238	−1.332	.001*	−1.148	.005*	−0.612	.229
Digit symbol	220	−7.654	.001*	−5.498	.019	−0.053	.986
Digit span forward	235	−0.556	.074	−0.354	.260	0.492	.194
Digit span backward	234	−1.101	<.001*	−0.833	.007*	0.172	.636
Trails A	232	24.855	<.001*	20.387	.002*	1.055	.897
Trails B	182	39.108	.031	28.348	.111	−5.607	.796

Note. MMSE = Mini-Mental Status Exam; CERAD = Consortium to Establish a Registry for Alzheimer's Disease; WL = Word List; LM = Logical Memory.
[a]WNHs were set as the reference category.
*Statistically significant to alpha level of .01.

https://doi.org/10.1017/S1355617711001688 Published online by Cambridge University Press

Table 4. Adjusted means before and after controlling for WTAR, with corresponding *p* values

	Adjusted means[a] (before WTAR)			Adjusted means[a] (after WTAR)		
	AA	WNH	*p* value	AA	WNH	*p* value
MMSE	19.1	21.8	.001*	21.1	21.3	.783
CERAD total	39.6	47.4	.0003*	45.4	45.9	.829
WL memory	4.3	4.6	.412	4.7	4.5	.573
WL delayed recall	0.9	1.5	.258	1.4	1.4	.834
WL recognition	13.8	16.0	<.0001*	14.5	15.8	.026
LM immediate	3.8	3.7	.884	4.5	3.5	.117
LM delay	0.9	1.3	.225	1.3	1.1	.588
Boston naming	14.6	20.1	<.0001*	17.9	19.3	.266
Category Fluency	9.3	10.2	.180	10.3	9.9	.575
Clock draw	5.3	4.6	.160	4.6	4.8	.688
Clock copy	3.2	2.9	.372	2.3	3.2	.063
Praxis construction	7.7	8.5	.023	8.5	8.3	.501
Praxis recall	1.4	2.6	.005*	1.8	2.5	.229
Digit symbol	20.3	25.8	.019	24.6	24.6	.985
Digit span forward	6.9	7.3	.260	7.7	7.2	.194
Digit span backward	4.0	4.9	.007*	4.8	4.7	.635
Trails A	86.7	66.4	.002*	71.3	70.3	.896
Trails B	240.6	212.2	.111	211.6	217.2	.795

Note. MMSE = Mini-Mental Status Exam; CERAD = Consortium to Establish a Registry for Alzheimer's Disease; WL = Word List; LM = Logical Memory; WTAR = Wechsler Test of Adult Reading.
[a]Adjusted for age, gender, and education.
*Statistically significant to alpha level of .01.

group differences in all of these tests were significantly attenuated even in this sample that was stratified by disease duration. Results for these variables are shown in Table 6.

Thus, the above results suggest that the observed WTAR effects are unlikely to be due to differences in disease severity in that significant attenuations of ethnoracial differences

Table 5. Linear regression models depicting ethnoracial differences in groups matched on functional severity

	AA-WNH coeff (w/o education or WTAR)[a]	*p* value	AA-WNH coeff (w/ education, w/o WTAR)[a]	*p* value	AA-WNH coeff (w/ education and WTAR)[a]	*p* value
MMSE	−2.61	<.0001*	0.198	.004*	0.057	.506
CERAD total	−0.246	<.001*	−0.199	.004*	0.073	.390
WL memory	−0.064	.379	−0.040	.5930	−0.030	.767
WL delayed recall	−0.147	.042	−0.137	.065	0.028	.769
WL recognition	−0.307	<.001*	−0.276	<.001*	−0.010	.919
LM immediate	−0.080	.286	−0.049	.519	0.099	.308
LM delay	0.014	.856	0.052	.498	0.222	.022
Boston naming	−0.336	<.001*	−0.294	<.001*	−0.021	.793
Category Fluency	−0.084	.238	−0.056	.440	0.112	.226
Clock draw	0.095	.203	0.058	.441	−0.139	.148
Clock copy	0.080	.283	0.025	.739	−0.252	.006
Praxis construction	−0.169	.020	−0.120	.098	0.130	.154
Praxis recall	−0.187	.012*	−0.160	.034	−0.024	.804
Digit symbol	−0.192	.012*	−0.140	.066	0.126	.198
Digit span forward	−0.218	.003*	−0.166	.024	0.068	.453
Digit span backward	−0.096	.195	−0.056	.449	0.096	.311
Trails A	0.209	.004*	0.173	.018	−0.081	.379
Trails B	1.605	.111	1.173	.243	−0.758	.450

Note. MMSE = Mini-Mental Status Exam; CERAD = Consortium to Establish a Registry for Alzheimer's disease; WL = Word List; LM = Logical Memory.
[a]WNHs were set as the reference category.
*Statistically significant to alpha level of .01.

https://doi.org/10.1017/S1355617711001688 Published online by Cambridge University Press

Table 6. Linear regression models depicting ethnoracial differences in groups stratified by disease duration*

	AA-WNH coeff (w/o education or WTAR)[a]	*p* value	AA-WNH coeff (w/ education, w/o WTAR)[a]	*p* value	AA-WNH coeff (w/ education and WTAR)[a]	*p* value
MMSE	−0.248	.001*	−0.176	.014*	−0.027	.746
CERAD total	−0.254	<.0001*	−0.196	.006*	−0.013	.870
WL memory	−0.013	.865	0.023	.766	0.073	.435
WL delayed recall	0.131	.085	−0.117	.131	0.011	.908
WL recognition	−0.300	<.0001*	−0.265	<.0001*	−0.139	.113
LM immediate	0.001	.989	0.039	.623	0.134	.151
LM delay	−0.086	.264	−0.061	.436	−0.073	.515
Boston naming	−0.385	<.0001*	−0.327	<.0001*	−0.142	.056
Category Fluency	−0.092	.217	−0.062	.413	0.049	.585
Clock draw	0.142	.067	0.109	.188	0.004	.987
Clock copy	0.071	.365	0.014	.856	0.148	.107
Praxis construction	0.176	.021	0.114	.126	0.034	.696
Praxis recall	−0.220	.004*	−0.193	.014*	−0.132	.154
Digit symbol	−0.134	.094	−0.084	.291	0.030	.757
Digit span forward	−0.049	.527	−0.009	.911	0.139	.127
Digit span backward	−0.167	.033	−0.119	.126	0.059	.511
Trails A	0.225	.003*	0.184	.014*	0.025	.774
Trails B	0.097	.264	0.083	.334	−0.041	.685

Note. MMSE = Mini-Mental Status Exam; CERAD = Consortium to Establish a Registry for Alzheimer's Disease; WL = Word List; LM = Logical Memory.
[a]WNHs were set as the reference category.
*Data only presented for subjects with short disease duration (<4 years).

persisted even in: a) groups that were matched on functional severity, and b) a sample that was stratified by disease duration. As such, these findings provide further support for the importance of accounting for reading ability when comparing neuropsychological performance between ethnoracial groups.

DISCUSSION

This study evaluated the extent to which reading ability, as an estimate of quality of education, accounts for the ethnoracial differences in cognitive performance between African Americans and WNHs. After controlling for the traditional demographic variables of age, sex, and years of formal education, African Americans scored significantly lower than WNHs on several of the neuropsychological tests. However, inclusion of WTAR as a proxy for quality of education eliminated all ethnoracial differences in performance on any of the tests. These results suggest that total years of schooling alone is an inadequate measure of the impact of education on cognitive performance in African Americans, and that differences in quality of education need to also be considered. More specifically, the number of years of formal education one has received may overestimate the reading level of African Americans compared to WNHs. This discrepancy may account for observed ethnoracial differences in cognitive performance that may have previously been attributed to biological or cultural differences between racial groups (Fillenbaum et al., 1997; Welsh et al., 1995; Wood et al., 2006).

The findings in this study are consistent with previous research conducted by Manly and colleagues (2002) in a community-based cohort of cognitively normal African American and non-Hispanic white elders. The authors used the Wide Range Achievement Test 3 (WRAT-3) Reading subtest (Wilkinson, 1993), as an estimate of quality of education, and found that adjusting for this variable attenuated differences in cognitive test performance between African Americans and WNH normal older adults matched on years of education (Manly et al., 2002). The current study extends these findings in a cohort of patients diagnosed with AD, and suggests that reading level is an important control factor in the evaluation of patients with suspected or established AD.

One hypothetical caveat to our results is that performance on the WTAR may have been confounded by the effects of dementia on reading ability. In other words, our finding that differences in WTAR score co-vary with differences in various measures of cognitive performance between African Americans and WNHs may in fact reflect that patients with more severe deficits in other cognitive domains are also poorer readers due to their illness rather than to differences in reading ability. Mitigating against this argument, prior evidence suggests that vocabulary and oral reading skills are relatively well preserved in patients with AD until late in the course of disease. For this reason, various investigators have used reading measures as a stable measure of premorbid performance in patients with cognitive impairment (Maddrey, Cullum, Weiner, & Filley, 1996; Paque & Warrington, 1995; Schmand, Geerlings, Jonker, & Lindeboom, 1998). Others, however, have debated whether reading performance remains

https://doi.org/10.1017/S1355617711001688 Published online by Cambridge University Press

stable in patients with AD (McFarlane, Welch, & Rodgers, 2006; O'Carroll et al., 1995).

To address these concerns, we controlled for severity of dementia in our analyses by using the FAQ, an assessment of functional severity. The addition of FAQ scores resulted in significant attenuation of ethnoracial differences in several of the neuropsychological test scores, indicating that ethnoracial differences in severity of dementia partially explain differences in test performance. However, the remaining differences in test scores between African Americans and WNHs were fully attenuated only after adjusting for WTAR. Similar results were obtained after matching racial groups on disease duration, another potential measure of disease severity. These findings suggest that differences in disease severity do not fully explain ethnoracial differences in reading scores and further supports the assumption that WTAR is a stable measure of premorbid functioning. Our results provide evidence that WTAR is a valid and useful control factor in the assessment of cognitive performance in a racially diverse and cognitively impaired population.

African Americans in our study sample were more significantly impaired than WNHs according to scores on the DSRS, CDR, and NPI-Q. However, it is interesting to note that controlling for quality of education attenuated all differences in test performance even without controlling for these differences in dementia severity. There are a couple of potential explanations for this counter-intuitive finding. First, it is possible that the presentation of AD differs between African Americans and WNHs such that African Americans exhibit more severe clinical symptoms compared to WNHs for any given level of cognitive impairment. Second, clinical rating scales may suffer from similar ethnoracial biases as neuropsychological test scores. For example, the CDR score is determined in part by the clinical assessment of cognition, which may bias against African Americans when not appropriately controlled for by quality of education. Either of these hypotheses may explain the apparent discrepancy between scores of dementia severity and neuropsychological performance.

A potential limitation of this investigation is that our cohort of patients may not be representative of the general population. Patients in this cohort are either self-referred, referred by their primary care providers, or recruited through outreach efforts in the West and North Philadelphia areas. Larger population-based studies will be needed to accurately characterize ethnoracial differences in performance on neuropsychological testing and the effect of WTAR score on attenuating those differences.

The effect of race on neuropsychological performance remains controversial (Fillenbaum, Heyman, Huber, Ganguli, & Unverzagt, 2001; Welsh et al., 1995). Number of years of education may inadequately estimate premorbid cognitive function for the purpose of standardizing cognitive assessments across ethnoracial groups. Though more research is required, the use of an oral reading assessment has the potential to reduce racial biases in the diagnosis and assessment of AD.

ACKNOWLEDGMENTS

There are no conflicts of interest for any of the authors. This work was supported by the American College of Neuropsychopharmacology Educational Outreach Fellowship (A.L.C.) as well as a K12 RR24151 grant (S.N.).

REFERENCES

Byrd, D.A., Touradji, P., Tang, X., & Manly, J.J. (2004). Cancellation test performance in African American, Hispanic, and White elderly. *Journal of the International Neuropsychological Society*, *10*, 401–411.

Chandler, M.J., Lacritz, L.H., Hynan, L.S., Barnard, H.D., Allen, G., Deschner, M., … Cullum, C.M. (2005). A total score for the CERAD neuropsychological battery. *Neurology*, *65*, 102–106.

Clark, C.M., & Ewbank, D.C. (1996). Performance of the dementia severity rating scale: A caregiver questionnaire for rating severity in Alzheimer disease. *Alzheimer Disease & Associated Disorders*, *10*, 31–39.

Demirovic, J., Prineas, R., Loewenstein, D., Bean, J., Duara, R., Sevush, S., & Szapocznik, J. (2003). Prevalence of dementia in three ethnic groups: The South Florida program on aging and health. *Annals of Epidemiology*, *13*, 472–478.

Fillenbaum, G.G., Heyman, A., Huber, M.S., Ganguli, M., & Unverzagt, F.W. (2001). Performance of elderly African American and White community residents on the CERAD Neuropsychological Battery. *Journal of the International Neuropsychological Society*, *7*, 502–509.

Fillenbaum, G.G., Huber, M., & Taussig, I.M. (1997). Performance of elderly White and African American community residents on the abbreviated CERAD Boston Naming Test. *Journal of Clinical and Experimental Neuropsychology*, *19*, 204–210.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (2004). *Mini mental state examination*. Lutz, FL: Psychological Assessment Resources.

IBM SPSS, Inc. (2009). *PASW statistics 18*. Chicago, IL: IBM SPSS, Inc.

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test*. Philadelphia, PA: Lea and Febiger.

Kaufer, D.I., Cummings, J.L., Ketchel, P., Smith, V., MacMillan, A., Shelley, T., … DeKosky, S.T. (2000). Validation of the NPI-Q, a brief clinical form of the Neuropsychiatric Inventory. *Journal of Neuropsychiatry and Clinical Neurosciences*, *12*, 233–239.

Maddrey, A.M., Cullum, C.M., Weiner, M.F., & Filley, C.M. (1996). Premorbid intelligence estimation and level of dementia in Alzheimer's disease. *Journal of the International Neuropsychological Society*, *2*, 551–555.

Manly, J.J., Jacobs, D., & Mayeux, R. (1999). Alzheimer's disease among different ethnic and racial groups. In R.D. Terry, R. Katzman, & K.L. Bick (Eds.), *Alzheimer's disease* (2nd ed., pp. 117–131). Philadelphia, PA: Lippincott Williams & Wilkins.

Manly, J.J., Jacobs, D.M., Sano, M., Bell, K., Merchant, C.A., Small, S.A., & Stern, Y. (1998). Cognitive test performance among nondemented elderly African Americans and whites. *Neurology*, *50*, 1238–1245.

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*, 341–348.

Manly, J.J., Touradji, P., Tang, X., & Stern, Y. (2003). Literacy and memory decline among ethnically diverse elders. *Journal of Clinical and Experimental Neuropsychology*, *25*, 680–690.

McFarlane, J., Welch, J., & Rodgers, J. (2006). Severity of Alzheimer's disease and effect on premorbid measures of intelligence. *British Journal of Clinical Psychology*, *45*, 453–463.

Morris, J.C. (1997). Clinical dementia rating: A reliable and valid diagnostic and staging measure for dementia of the Alzheimer type. *International Psychogeriatrics*, *9*, 173–176.

Morris, J.C., Heyman, A., Mohs, R.C., Hughes, J.P., van Belle, G., Fillenbaum, G., … Clark, C. (1989). The consortium to establish a registry for Alzheimer's disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer''s disease. *Neurology*, *39*, 1159–1165.

O'Carroll, R.E., Prentice, N., Murray, C., van Beck, M., Ebmeier, K.P., & Goodwin, G.M. (1995). Further evidence that reading ability is not preserved in Alzheimer's disease. *British Journal of Psychiatry*, *167*, 659–662.

Paque, L., & Warrington, E.K. (1995). A longitudinal study of reading ability in patients suffering from dementia. *Journal of the International Neuropsychological Society*, *1*, 517–524.

Pedraza, O., & Mungas, D. (2008). Measurement in cross-cultural neuropsychology. *Neuropsychology Review*, *18*, 184–193.

Petersen, R.C., Smith, G.E., Ivnik, R.J., Tangalos, E.G., Schaid, D.J., Thibodeau, S.N., … Kurland, L.T. (1995). Apolipoprotein E status as a predictor of the development of Alzheimer's disease in memory-impaired individuals. *Journal of the American Medical Association*, *273*, 1274–1278.

Pfeffer, R.I., Kurosaki, T.T., Harrah, C.H., Chance, J.M., & Filos, S. (1982). Measurement of functional activities in older adults in the community. *Journal of Gerontology*, *37*, 323–329.

Reitan, R., & Wolfson, D. (1993). *The Halstead-Reitan Neuropsychological Test Battery: Theory and clinical interpretation* (2nd ed.,). Tucson, AZ: Neuropsychology Press.

Schmand, B., Geerlings, M.I., Jonker, C., & Lindeboom, J. (1998). Reading ability as an estimator of premorbid intelligence: Does it remain stable in emergent dementia? *Journal of Clinical and Experimental Neuropsychology*, *20*, 42–51.

Stephenson, J. (2001). Racial barriers may hamper diagnosis, care of patients with Alzheimer disease. *Journal of the American Medical Association*, *286*, 779–780.

Tang, M.X., Cross, P., Andrews, H., Jacobs, D.M., Small, S., Bell, K., … Mayeux, R. (2001). Incidence of AD in African-Americans, Caribbean Hispanics, and Caucasians in northern Manhattan. *Neurology*, *56*, 49–56.

Tombaugh, T.N., & McIntyre, N.J. (1992). The mini-mental state examination: A comprehensive review. *Journal of the American Geriatrics Society*, *40*, 922–935.

Wechsler, D. (1987a). *Wechsler memory scale-revised manual*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1987b). *Wechsler adult intelligence scale—Revised*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997a). *Wechsler adult intelligence scale—Third edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (1997b). *Wechsler memory scale—Third edition*. San Antonio, TX: The Psychological Corporation.

Wechsler, D. (2001). *Wechsler test of adult reading*. San Antonio, TX: The Psychological Corporation.

Weintraub, S., Salmon, D., Mercaldo, N., Ferris, S., Graff-Radford, N.R., Chui, H., … Morris, J.C. (2009). The Alzheimer's Disease Centers' Uniform Data Set (UDS): The Neuropsychological Test Battery. *Alzheimer Disease & Associated Disorders*, *23*, 91–101.

Welsh, K.A., Fillenbaum, G., Wilkinson, W., Heyman, A., Mohs, R.C., Stern, Y., … Beekly, D. (1995). Neuropsychological test performance in African-American and white patients with Alzheimer's disease. *Neurology*, *45*, 2207–2211.

Whitfield, K.E., & Wiggins, S. (2003). The impact of desegregation on cognition among older African Americans. *Journal of African American Psychology*, *29*, 275–291.

Wilkinson, G.S. (1993). *Wide range achievement test 3*. Wilmington, DE: Wide Range, Inc.

Wood, R.Y., Giuliano, K.K., Bignell, C.U., & Pritham, W.W. (2006). Assessing cognitive ability in research: Use of MMSE with minority populations and elderly adults with low education levels. *Journal of Gerontologic Nursing*, *32*, 45–54.

Xie, S.X., Ewbank, D.C., Chittams, J., Karlawish, J.H., Arnold, S.E., & Clark, C.M. (2009). Rate of decline in Alzheimer disease measured by a Dementia Severity Rating Scale. *Alzheimer Disease & Associated Disorders*, *23*, 268–274.

Yesavage, J.A., Brink, T.L., Rose, T.L., Lum, O., Huang, V., Adey, M., & Leirer, V.O. (1982). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, *17*, 37–49.