

Reading ability mediates the relationship between education and executive function tasks

AMANDA SCHAFER JOHNSON, LISA J. FLICKER, AND PETER A. LICHTENBERG

Institute of Gerontology, Wayne State University, Detroit, Michigan

(RECEIVED February 14, 2005; FINAL REVISION October 2, 2005; ACCEPTED October 3, 2005)

Abstract

Neuropsychological test results are affected by multiple factors, but usually age and education are the only variables by which norms are stratified. Some authors have questioned whether these variables alone are sufficient (e.g., Marcopulos et al., 1997; Manly et al., 2002), since such norms have led to problems, such as poor specificity for African Americans on dementia screening devices (Fillenbaum et al., 1990). Recent research has shown that reading ability, a measure of educational quality, attenuated racial differences in test performance (Manly et al., 2002). We specifically examined whether reading ability would account for a greater amount of variance than education in executive function tests in a population traditionally subject to poor educational quality. Results determined that reading ability accounted for a significantly greater amount of variance than years of education for Letter-Number Sequencing, Similarities, COWA, Trail Making Test, and Coloured Progressive Matrices. Reading ability was found to significantly mediate the relationship between each of these tests and education. Animal naming appears to be least affected by educational quality or quantity. These findings hold implications for the interpretation of neuropsychological test results, especially in those exposed to substandard educational quality, and for the way that test norms are constructed. (*JINS*, 2006, *12*, 64–71.)

Keywords: Literacy, Dementia, Norms, Educational quality, African Americans, Aged

INTRODUCTION

It is readily apparent that there are a variety of factors that impact neuropsychological test performance. As a result of this, clinical neuropsychology frequently stratifies norms by factors such as age, or less frequently, by both age and education. However, race is another factor that has repeatedly demonstrated a significant relationship with test scores (e.g., Reynolds et al., 1987; Welsh et al., 1995; Marcopulos et al., 1997). Since there is overwhelming evidence that test scores are affected by more than simply age and years of education, it is apparent that using these two covariates alone is unlikely to yield the most universal norms.

Poor norm development has shown deleterious effects for minority populations. For example, Fillenbaum and colleagues (1990) found six dementia screening devices to have significantly poorer specificity for African Americans com-

pared to whites. By definition, poor specificity implies false positive results and, in this case, an overestimation of cognitive impairment. Others have noted that based on their neuropsychological test scores alone, African Americans would be incorrectly classified as demented with a much greater frequency than whites (Fillenbaum et al., 1990; Gurland et al., 1992) and there have been warnings about the danger of misdiagnosing African Americans with dementia (Froehlich et al., 2001; Mast et al., 2001; Miles 2001).

In response to this overestimation of cognitive impairment there have been efforts to reconsider the way norms are currently established. Marcopulos and colleagues (1997) have suggested that covarying test scores by IQ along with age and education will lead to more accurate data interpretation, especially for individuals for whom education does not reflect cognitive ability. Manly and colleagues (2002) attempted to find a covariate that better helped to explain racial differences found in test performance. Rather than using IQ or educational quantity, they attempted to quantify educational quality. Using reading ability as a measure of educational quality, they found that many of the interracial differences became nonsignificant.

Reprint requests to: Amanda Schafer Johnson, M.A., Institute of Gerontology, Wayne State University, 87 E. Ferry Street, Detroit, MI 48202. E-mail: a.schafer@wayne.edu

It has been previously discussed by researchers that when years of education are held constant, cognitive outcome can vary substantially. For example, with reference to making corrections to normative values from one geographic region to another, Gurland and colleagues (1992) noted, “implausible is the implication in these adjustment formulae that a certain number of years of education in one school system is equivalent [in psychometric influence] to the same number of years in a different educational system” (p. 108). There are also reported differences in educational quality, such as student/teacher ratios, per pupil expenditures, access to facilities, and the like that have been shown to correlate significantly with standardized achievement scores (for a review, see Manly et al., 2002) and by definition, educational outcomes. When one considers the influence of these idiosyncratic educational factors along with variance due to heritability and parental encouragement of education, it appears even more unlikely that equivalent years of education will necessarily yield equivalent cognitive outcomes.

Because other research studies have found success using reading ability to explain excess variance in test scores beyond age and years of education, this study aimed to examine the impact that reading ability has specifically on executive function (EF) tests. As the participants in this study were specifically older adults, EF tests were chosen due to their early decline resulting from frontal lobe atrophy in normal aging (e.g., Double et al., 1996).

In light of these findings, the following hypotheses will be tested by this study: (1) Reading ability will account for significant variance in EF tests above and beyond that accounted for by education. (2) Reading ability will statistically mediate the relationship between education and EF tests.

METHOD

Research Participants

One hundred participants were recruited from the city of Detroit; all were between the ages of 59 and 95. Participants were recruited from various independent living centers and community centers, through the distribution of flyers and by word of mouth. Participant education ranged from 5 to 20 years. Ninety-three percent were African American and 81.4% were female. This reflects a slightly greater proportion of females than in the Detroit older adult cohort, which is 70% female (Chaplewski, 2002), but is not inconsistent with other studies of older, African American adults (e.g., Albert & Teresi, 1999; Manly et al., 2002).

No special inclusion or exclusion criteria were enforced, with the exception that participants must be above age 59 and native English speakers. Willing participants who were sufficiently cognizant to understand and comply with instructions were admitted to the study. Because this study was interested in the influence reading ability had on all people,

participants were not excluded based on cognition. Implications for this are in the Discussion section. Twenty dollars cash was given to each participant as compensation for his or her time.

Of the 100 participants tested, 97 were administered the WRAT-3 Reading scale. Time constraint was the only reason that 3 of the participants were not administered the WRAT-3. Because this scale was pivotal to the purpose of this study, only the 97 that were administered the measure were included in analyses.

Apparatus

Wide Range Achievement Test-3 (WRAT-3): Reading (Wilkinson, 1993)

This is a test of word familiarity and sight reading ability involving the pronunciation of a series of 15 letters of the alphabet and 42 increasingly difficult words. A maximum score of 57 is possible. As reported in Spreen and Strauss (1998), it has been useful in estimating premorbid intelligence, and moderately relates to WAIS-R IQ scores (Griffin et al., 2002).

Verbal Fluency (Benton & Hamsher, 1976)

This series of brief tests involves rapidly naming items that fit a description provided by the administrator. This test utilized both Controlled Oral Word Association (COWA-Letters /C/, /F/, /L/) and category fluency (Animal Naming). A principal components analysis (Marcie et al., 1993) found verbal fluency to load onto an “abstract mental operations” factor, along with other tasks such as mental calculations and Digit Span. As reviewed by Lezak (1995), frontal lobe lesions are consistently linked to lowered verbal fluency output, although there is some evidence that the task is mediated by the temporal lobes as well as the frontal.

Wechsler Adult Intelligence Scale-III (WAIS-III): Similarities (Wechsler, 1997)

This task involves the presentation of 19 increasingly disparate pairs of words (e.g., dog and lion). Participants are asked to state how the two words are alike. A maximum of score of 33 is possible. The number of errors on this task has been implicated as an indication of dysexecutive syndrome seen in dementing disorders (Giovannetti et al., 2001). The Similarities subtest has been used in many studies as a measure of executive function (see Giovannetti et al., 2001).

WAIS-III: Letter-Number Sequencing (Wechsler, 1997)

Participants are read a series of numbers and letters, and are instructed to mentally rearrange the digits so that the numbers are said in order, followed by the letters said in alphabetical order. Participants begin by manipulating 2 digits, and work up to 8 digits or until they meet discontinu-

teria. Maximum score is 21. Factor analyses have consistently shown this subtest to be among those thought to tap working memory (Wechsler, 1997; Ryan & Paolo, 2001), a facility commonly associated with prefrontal, executive function (Loring, 1999).

Coloured Progressive Matrices (Raven, 1965)

For this test, participants are shown 36 individually presented visual patterns and asked to identify the missing part of the target design from a set of choices. The maximum score is 36. CPM has been implicated as a means of tapping fluid intelligence; as reported by Mills and colleagues (1993), it has a strong correlation with the WAIS Block Design subtest, which is widely regarded as being a good measure of fluid intelligence. Fluid intelligence is the ability to reason and problem solve in the absence of familiar solutions (Lezak, 1995); this is very similar to the concept of executive ability.

Trail Making Test (Reitan & Wolfson, 1985)

For part A of this task, participants are asked to connect numbered circles as rapidly as possible. For part B, participants are asked to connect circles containing alternating numbers and letters (A-1-B-2) as quickly as possible. Time to completion is noted; a maximum of 300 seconds is permitted. As reviewed by Spreen and Strauss (1998), factor analyses have shown Trails B to load onto factors such as “rapid visual search,” “visuospatial sequencing,” “focused mental processing speed,” and “cognitive set shifting,” all of which are related to the concept of executive ability.

CLOX (Royall et al., 1998)

This clock-drawing test instructs participants to draw two clocks. For the first (CLOX 1), they are simply given oral instructions to “Draw a clock. Set the hands and the numbers on the face so it reads 1:45.” The second clock (CLOX 2) is copied from an example drawn by the examiner. Both clocks are scored on 15 criteria outlined by Royall and colleagues (1998), and a maximum score of 15 on each clock is possible. These criteria include the similarity of the drawing to a clock, the presence of an exterior circle, the circle being larger than one inch, numbers being within the circle, numbers being evenly spaced, only the presence of numbers 1–12, the numbers placed in correct numeric order, the numbers expressed as Arabic numerals, the placement of 3, 6, 9, and 12 first, the hour hand between the 1 and 2, the minute hand being longer than the hour hand, the hands being represented as arrows, and the absence of such items as the drawing of a hand or a face, intrusion from a circle on the other side of the page, the writing of “1:45,” hands pointing to 4 or 5 o’clock, nor any words, letters, or pictures. The Royall CLOX test has been shown to explain significant variance of an executive function interview (EXIT 25), and was the only method of clock drawing that did so when compared to other methods (Royall et al., 1999). The

first clock drawing has been purported to be more dependent on executive ability than the second; as such, it has also been found to be more highly correlated to the same executive ability interview, as would be expected (Royall et al., 1998).

Procedure

This study is part of a larger data collection effort that also includes measures of health, stressors, and perceived social support; the whole battery took 2–2.5 hours to administer. Administration took place in various places in the community: offices of community centers, assisted living centers, and participant’s homes for those who felt comfortable with this. After participants were recruited (see Participants section), each was administered the previously mentioned battery of tests. The battery was no less than 2 hours, and longer for those who were slow or garrulous. Consequently, the battery was cut short for some participants; when measures were not administered, examiners attempted to randomize which variables these were on a case-by-case basis. There were missing data for each variable except for age and education; out of a total of 100 original participants, the total that were administered each test is listed in Table 2.

Prior to conducting the analyses, skewness was examined and corrected. To examine the relationship between reading and executive ability, a series of 9 hierarchical multiple regression analyses were conducted. Raw scores on each of the executive ability tests were dependent variables for each individual regression. Predictor variables included age, years of education, and raw scores on the WRAT-III Reading test. Specifically, age and education were predictor variables in Block 1, and WRAT-3 Reading score was the predictor variable added in Block 2. This design permitted an examination of the effect of reading ability on test scores after the effects of age and education had already been taken into account. To preserve the integrity of the alpha level and reduce the possibility of Type I or Type II errors, a moderately conservative alpha of .01 was considered significant. Probabilities of .05 are noted in tables to illustrate trends.

Missing data were examined by dummy-coding each variable and making simple *t* test comparisons between those who had been administered each variable and those who had not. None of the variables met significance, indicating that those with and without missing data on each variable did not significantly differ with regard to demographic characteristics or performance on other tests. To further address the missing data, analyses were first run using pairwise deletions and then using listwise deletions to determine whether missing cases influenced results; as seen in the Results section, they did not.

RESULTS

Characteristics of the overall sample and their test results are reported in Tables 1 and 2, respectively. It was noted

Table 1. Demographics and health characteristics of the sample

	<i>n</i>	Mean/Percentage	<i>SD</i>
Age	97	72.34	7.89
Education	97	12.34	3.12
Gender			
Female	79	81.4%	
Male	18	18.6%	
Ethnicity			
African American	91	93.8%	
European American	4	4.1%	
Other	2	2.1%	
Medical history significant for			
Stroke	21	21.6%	
TIA	10	10.3	
Hypertension	67	69.1%	
Diabetes	27	27.8%	
Head injury w/LOC	5	5.2%	
Heart Attack	14	14.4%	
Parkinson's disease	2	2.1%	
Huntington's disease	0	0%	
Multiple sclerosis	1	1.0%	

that a sizable proportion (21.6%) of the participants had previously experienced strokes. Because of this, the analyses were conducted with and without those with stroke history. As no differences in data interpretation were found, participants with strokes remained in this study.

CLOX 1, CLOX 2, and Trail Making Test A were significantly skewed. Both CLOX 1 and CLOX 2 had a single low outlying variable; the deletion of these variables eliminated significant negative skew. Significant positive skewness remained for Trail Making Test A even after eliminating its outlier; consequently, the square root of this variable was taken to transform the variable into one with a normal distribution. Listwise and pairwise regressions were both examined in the following analyses, and no difference was found between them; subsequently, the more conservative

Table 2. Sample test results

	<i>n</i>	Mean	<i>SD</i>
WRAT Reading	97	39.87	7.03
Letter-Number	94	6.64	2.82
COWA	95	26.73	11.22
Animal Naming	93	13.41	4.34
CPM	95	23.00	6.21
CLOX 1	90	10.38	2.42
CLOX 2	90	12.28	2.14
Similarities	89	13.98	7.00
Trails A	89	79.94	44.09
Trails B	88	180.16	84.03

Note. Letter-Number = WAIS-III Letter-Number Sequencing, COWA = Controlled Oral Word Association Test, CPM = Coloured Progressive Matrices, Similarities = WAIS-III Similarities.

listwise deletions are reported in the data that follows. As a result of using listwise deletions for each regression that was conducted, individuals having missing data for the criterion variable were eliminated from that particular analysis. Because there were different regressions conducted for each EF variable, the specific individuals that were eliminated varied per analysis.

Using the hierarchical regression analyses with pairwise deletions, R^2 Change values were significant at Block 1 for all executive function variables, indicating that all tests were suspect to the effects of age and/or education. R^2 Change was significant in Block 2 for WAIS-III Letter-Number Sequencing, COWA, Coloured Progressive Matrices, WAIS-III Similarities, and Trail Making Test B, indicating that for these 5 tests there was additional variance to be explained by reading ability even after variance due to age and education had been partialled out. See Table 3.

The beta values in Table 4 reveal that after the addition of reading ability in Block 2, correlations between test scores and years of education were reduced to nonsignificance. Since the relationship between years of education and executive function tests was reduced to nonsignificance for several of these variables when reading ability was entered into the regression, it is thought that reading ability may mediate their relationship. Reading ability's significance as a mediator was tested via the Sobel (1988) method, as outlined by Holmbeck (2002). This method examines the relationship between the predictor and criterion variables and determines whether any drop in effect size resulting from the addition of the mediator variable is statistically significant. It was not expected that reading ability would act as a mediator between years of education and test score for any of the EF tests that were not significantly influenced by reading ability—by definition this would counter the meaning of mediation. Indeed, none was significant. The relationships between the remaining EF tests and education were dropped to nonsignificance upon the entry of WRAT-3 Reading score; each of these decreases was shown to be statistically significant at $p < .05$. Z-scores for the indirect effects and the percent of the relationship mediated by reading ability are illustrated in Table 5.

Lastly, an analysis was conducted to determine how well the reported educational attainment of our participants corresponded with their demonstrated reading ability. Education and reading ability scores were correlated at $r = .58$ ($p < .01$). Of our participants, 73.19% ($n = 71$) read at a grade level below their stated educational attainment, 19.59% ($n = 19$) read at a grade level approximately equivalent to their stated educational attainment, and 7.22% ($n = 7$) read at a grade level above their educational attainment. See Figure 1 for a distribution of how these participants' reading levels differ from their educational attainment. People who had between 9 and 12 years of education were considered to be reading at grade level if they read at a "high school" equivalency. Likewise, people with 13 years of education or more were considered to read at grade level if they read at "post high school" equivalency. These truncations

Table 3. *R* square change values

	LN (<i>n</i> = 94)	COWA (<i>n</i> = 95)	ANI (<i>n</i> = 93)	CPM (<i>n</i> = 95)	CLX1 (<i>n</i> = 89)	CLX2 (<i>n</i> = 89)	SIM (<i>n</i> = 89)	TMTA (<i>n</i> = 89)	TMTB (<i>n</i> = 88)
Block 1 Education and age	.14**	.15**	.23**	.21**	.12**	.07*	.32**	.20**	.21**
Block 2: Reading	.11**	.16**	.02	.06**	.00	.01	.13**	.04*	.13**

Note. LN = WAIS-III Letter-Number Sequencing, COWA = Controlled Oral Word Association Test, ANI = Animal Naming, CPM = Coloured Progressive Matrices, CLX1 = Royall's CLOX 1, CLX2 = Royall's CLOX 2, SIM = WAIS-III Similarities, TMTA = Trail Making Test A (transformed), TMTB = Trail Making Test B.

*Trend at $p < .05$ level (2-tailed).

**Significant at $p < .01$ level (2-tailed).

are necessary, as the WRAT-3 manual does not individuate into specific reading-level equivalency past eighth grade but compiles scores into the aforementioned “high school” and “post high school” categories. Because of these truncations in reading levels, considering absolute differences between scores is cautioned against.

DISCUSSION

This study determined that reading ability—a purported measure of educational quality—accounted for significant variance in several executive function tests above and beyond the variance explained by years of education. When reading ability was added in a hierarchical regression, the relationships between years of education and most EF variables (namely WAIS-III Letter-Number Sequencing and Similarities, COWA, Coloured Progressive Matrices, and the Trail Making Test) were no longer significant. This was consistent with the prediction that reading ability would have a greater influence on test performance and would explain more variance than simple educational quantity. Manly and colleagues have done a great deal of research in this area; the present study, however, demonstrates this effect with a larger variety of tests specific to executive abilities. It also

contributed to the literature through its second aim—its statistical demonstration that the relationship between test performance and years of education is significantly mediated by a measure of educational quality. For one test (COWA), as much as 98.81% of its relationship with years of education was explained by reading ability.

While reading ability did account for significant variance on 6 tests, it did not for 3 others. Specifically, Animal Naming, CLOX 1, and CLOX 2 were not related to reading ability. It may be that the concepts behind Animal Naming and CLOX are simpler than that of other tests. If they are conceptually more easy to understand than the other tests and do not require much capacity for abstraction, the quality of one's education may not make a difference. Alternatively, it seems possible that drawing a clock (CLOX) and listing species of animals (Animal Naming) are more familiar and less anxiety-provoking than other tasks.

Although there are some past studies that used reading ability as a predictor of test performance, these are relatively limited thus far. Albert and Teresi (1999) found reading ability to be a significant, independent predictor of Mini-Mental State Exam (MMSE) performance even when years of education were included in the equation. As mentioned earlier, Manly and colleagues (2002) found that many dif-

Table 4. β values, Blocks 1 and 2

	Block 1		Block 2		WRAT-3
	Age	Education	Age	Education	
LN	-.19	.29**	-.18	.04	.42**
COWA	-.19	.32**	-.16	.01	.51**
Animals	-.46**	.09	-.45**	.01	.16
CPM	-.27**	.32**	-.26**	.15	.30**
CLOX1	-.17	.27**	-.17	.27*	.00
CLOX2	-.08*	.01	-.08	.15	.14
Similarities	-.32**	.40**	-.29**	.14	.45**
TMT-A	.30**	-.31**	.28**	-.17	-.23*
TMT-B	.34**	-.27**	.31**	-.01	-.44**

Note. See list of abbreviations in Table 3 footnote. *Trend at $p < .05$ level (2-tailed).

**Significant at $p < .01$ level (2-tailed).

Table 5. Reading ability as a significant mediator

Executive function test	<i>z</i> value	Percent of relationship between test and education accounted for by reading ability
Letter-Number Sequencing	3.24*	79.62
COWA	4.02*	98.81
Animal Naming	1.37	ns
Coloured Progressive Matrices	2.52*	47.33
CLOX 1	0.23	ns
CLOX 2	1.15	ns
Similarities	3.87*	64.50
Trail Making Test A (transformed)	-2.19*	46.43
Trail Making Test B	-3.61*	95.97

*Significant at $p < .05$ level, ns = not significant.

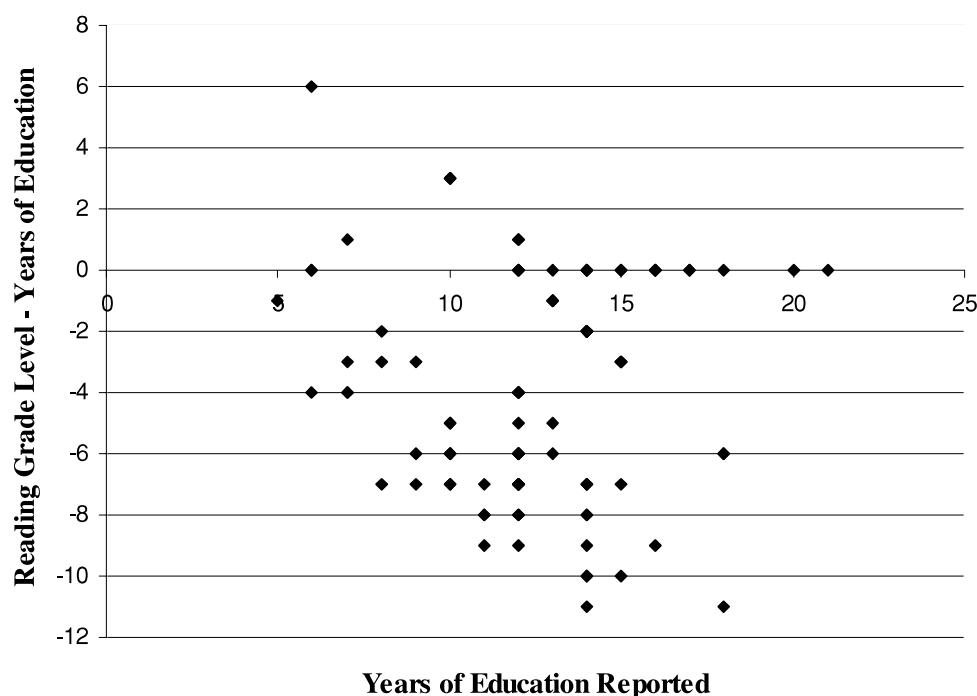


Fig. 1. Difference between reading grade level and educational attainment plotted as a function of educational attainment.

ferences between racial groups on neuropsychological tests were reduced to nonsignificance when reading score was introduced as a covariate. This included memory tests, EF tests (specifically, WAIS-R Similarities and DRS Identities and Oddities), a visuospatial matching task, and language tasks. In a 2004 follow-up study, this group determined that reading ability was still the best predictor of test performance even after a measure of acculturation was added to their regression. Along with the Boston Naming Test, COWA and Similarities were the variables most affected by reading ability.

The implications of this study are numerous. It demonstrated that there is a great deal of variance in cognitive ability even when educational attainment is held constant. This is illustrated well by Figure 1. For this particular cohort, it was found that reading ability was much lower than expected based on the participant's reported educational attainment, and it seems likely that other cognitive abilities (e.g., mathematical ability, reading comprehension) would be below grade-equivalency. Other studies have also found this to be the case: Using a sample of African American participants, Albert and Teresi (1999) found approximately 50% read at a level lower than grade-equivalency, and Manly and colleagues (2002) found this for 33% of its African American participants. The present study found that 73% of its participants read at a level lower than expected based on educational attainment. Thus, although many studies assume approximate equivalence in educational outcomes between people having equivalent educational quantity, several studies including the present have demonstrated that this is often not the case. Considering that educational quality (operationalized here by reading ability) signifi-

cantly affects test performance, variance among people with the same amount of education is likely to impact test interpretation.

There are obvious clinical implications of these results, as well. If individuals score poorly on such measures because the measures are heavily dependent on reading ability, their performance may erroneously be considered a pathological sign. Although attempts to control for such issues have been made by stratifying some norms by educational level, this and other studies reviewed here demonstrate the deficiencies of this approach because educational attainment often does not reflect underlying cognitive abilities. For example, a person reporting twelve years of education may only have learned to read at a sixth grade level, and his or her other cognitive abilities may also be consistent with that of someone having six years of education.

This begs the question of whether there are tests that are appropriate to use with persons of poor reading ability and/or persons who have received an overall substandard educational experience. This study determined that category fluency and a clock drawing test were not influenced by reading ability. Because clinicians are likely to continue using tests that have a significant relationship with educational quality, it would behoove them to use reading level to estimate premorbid intelligence. They would be able to make more accurate interpretations by taking the reading level of the patient into consideration.

Finally, there are implications for norm development. There are substantial problems with current norm development as demonstrated by both the vast disparity between actual ability and expectations based on educational attainment, and the interethnic differences resulting from the use

of current norms. As reading ability explains a greater amount of the variance in test scores compared to educational attainment (and largely mediates the relationship between years of education and test scores), designing norms stratified in part by reading ability would lead to more accurate test interpretation. This would be easy to accomplish, as sight reading tests such as WRAT-3 Reading are quick and painless to administer, making it simple for both researchers to include them in norm development and clinicians to include them for interpretation.

Understanding educational quality seems to be an especially important issue for African Americans and potentially for other ethnic minority groups, as well. In their 2002 study, Manly and colleagues found that only 7% of their European American participants read below stated grade level compared to 33% of African Americans. Ours, too, was an older cohort, many of whom anecdotally related having been raised in the southern U.S. before moving to Detroit. As such, many older African Americans were educated in the “separate but equal” conditions prior to the *Brown vs. the Board of Education* ruling. The quality of their educational experience was certainly poorer than that of their white peers, as can be seen by the high student-to-teacher ratios, low per student expenditures, less access to facilities, fewer hours in the school week, and poorer teacher quality compared to predominantly white schools (U.S. Department of Health, Education, and Welfare, 1966). This substandard educational experience may also be the reason that such a large percentage of African Americans are reading below grade equivalency in several samples. This phenomenon is certainly not limited to African Americans and other minority groups, however. Note that 7% of the white sample in Manly and colleagues’ (2002) study also read at expected grade equivalency. The base rate of any education-ability discrepancy will vary based on quality of the educational experiences of any population.

There are limitations to this study. Participant selection is one of these, as we accepted participants using relatively few exclusion criteria. Therefore, there were participants with histories of stroke, transient ischemic attack, head injury, and probably MCI or early dementia. Based on previous findings, reading ability scores should hold for all of these groups (Johnstone & Wilhelm, 1996), although the EF scores of some may have been negatively affected by pathology. As these findings may not be universal across the lifespan, another limitation of this study was that it included only older adults.

A final caveat about these research findings: While several researchers have connected reading ability to educational quality (e.g., Greenwald et al., 1996; Wilkinson, 1993; Manly et al., 2002), reading ability may be correlated to all of these measures not because of its properties as a measure of educational quality but because of the relationship it has with *g*. Reading ability is highly correlated with IQ scores. Griffin et al. (2002) found a correlation between WAIS-R FSIQ and WRAT-3 reading of $r = .63$ ($p < .001$), while Wilkinson (1993) reported a correlation of .53. Since our

reading measure correlates with *g* and *g* correlates well with most standardized tests (e.g., Diaz-Asper et al., 2004), it is entirely possible that a measure of reading ability predicts well to all cognitive tests because of the issue of intercorrelation. Even if it is no more than a proxy for *g*, reading ability is still quick and easy to ascertain, and unlike most other cognitive measures is considered a “hold” test in that it does not endure clinically significant declines early in most neurocognitive disorders, including mild to moderate dementia (Johnstone & Wilhelm, 1996; Schmand et al., 1998).

Reading level is not the panacea for poor norms. Through its use we can make more accurate interpretations of test data. Although covarying by reading ability attenuated significant differences between whites and African Americans on neuropsychological tests, racial disparities still existed (Manly et al., 2002). Further research ought to be directed at uncovering which covariates further narrow this racial gap.

ACKNOWLEDGMENTS

We are grateful to our participants in the Detroit community and others who helped connect us with participants, including staff members at Joseph Walker Williams Community Center, Hannan House, Arnold Home, and Brush Park Manor. The professional and technical support of Coletta Nelson-Thomas, Annmarie Cano, Scott Moffat, Paul Cernin, and Thomas Jankowski is also appreciated. Financial support for this research was provided by internal funds of the Wayne State University Institute of Gerontology.

REFERENCES

- Albert, S.M. & Teresi, J.A. (1999). Reading ability, education, and cognitive status assessment among older adults in Harlem, New York City. *American Journal of Public Health, 89*, 95–97.
- Benton, A.L. & Hamsher, K. (1976). *Multilingual Aphasia Examination*. Iowa City: University of Iowa.
- Chaplewski, E. (2002). *Facing the future: City of Detroit needs assessment of older adults*. Detroit, MI: Department of Senior Citizens.
- Diaz-Asper, C.M., Schretlen, D.J., & Pearlson, G.D. (2004). How well does IQ predict neuropsychological test performance in normal adults? *Journal of the International Neuropsychological Society, 10*, 82–90.
- Double, K.L., Halliday, G.M., Kril, J.J., Harasty, J.A., Cullen, K., Brooks, W.S., Creasey, H., & Broe, G.A. (1996). Topography of brain atrophy during normal aging and Alzheimer’s Disease. *Neurobiology of Aging, 17*, 513–521.
- Fillenbaum, G., Heyman, A., William, K., Prosnitz, B., & Burchett, B. (1990). Sensitivity and specificity of standardized screens of cognitive impairment and dementia among elderly black and white community residents. *Journal of Clinical Epidemiology, 43*, 651–660.
- Froehlich, T.E., Bogardus, S.T., & Inouye, S.K. (2001). Dementia and race: Are there differences between African Americans and Caucasians? *Journal of the American Geriatrics Society, 49*, 477–484.
- Giovannetti, T., Lamar, M., Cloud, B.S., Swenson, R., Fein, D., Kaplan, E., & Libon, D.J. (2001). Different underlying mech-

- anisms for deficits in concept formation in dementia. *Archives of Clinical Neuropsychology*, *16*, 547–560.
- Greenwald, R., Hedges, L.V., & Laine, R.D. (1996). The effect of school resources on student achievement. *Review of Educational Research*, *66*, 361–396.
- Griffin, S.L., Mindt, M.R., Rankin, E.J., Ritchie, A.J., & Scott, J.G. (2002). Estimating premorbid intelligence: Comparison of traditional and contemporary methods across the intelligence continuum. *Archives of Clinical Neuropsychology*, *17*, 497–507.
- Gurland, B.J., Wilder, D.E., Cross, P., Teresi, J., & Barrett, V.W. (1992). Screening scales for dementia: Toward reconciliation of conflicting cross-cultural findings. *International Journal of Geriatric Psychiatry*, *7*, 105–113.
- Holmbeck, G.N. (2002). Post-hoc probing of significant moderational and mediational effects in studies of pediatric populations. *Journal of Pediatric Psychology*, *27*, 87–96.
- Loring, D.W. (Ed.). (1999). *INS dictionary of neuropsychology*. New York: Oxford University Press.
- Johnstone, B. & Wilhelm, K.L. (1996). The longitudinal stability of the WRAT-R Reading subtest: Is it an appropriate estimate of premorbid intelligence? *Journal of the International Neuropsychological Society*, *2*, 282–285.
- Lezak, M.D. (1995). *Neuropsychological assessment* (3rd ed.). New York: Oxford University Press.
- Manly, J.J., Jacobs, D.M., Touradji, P., Small, S.A., & Stern, Y. (2002). Reading level attenuates difference in neuropsychological test performance between African American and White elders. *Journal of the International Neuropsychological Society*, *8*, 341–348.
- Manly, J.J., Byrd, D.A., & Touradji, P. (2004). Acculturation, reading level, and neuropsychological test performance among African American elders. *Applied Neuropsychology*, *11*, 37–46.
- Marcie, P., Roudier, M., Goldblum, M., & Boller, F. (1993). Principal component analysis of language performances in Alzheimer's disease. *Journal of Communication Disorders*, *26*, 53–63.
- Marcopulos, B.A., McLain, C.A., & Giuliano, A.J. (1997). Cognitive impairment or inadequate norms? A study of healthy, rural, older adults with limited education. *The Clinical Neuro-psychologist*, *11*, 111–131.
- Mast, B.T., Fitzgerald, J., Steinberg, J., MacNeill, S.E., & Lichtenberg, P.A. (2001). Effective screening for Alzheimer's disease among older African Americans. *The Clinical Neuro-psychologist*, *15*, 196–202.
- Miles, T.P. (2001). Dementia, race, and education: A cautionary note for clinicians and researchers. *Journal of the American Geriatrics Society*, *49*, 490.
- Mills, C.J., Ablard, K.E., & Brody, L.E. (1993). The Raven's Progressive Matrices: Its usefulness for identifying gifted/talented students. *Roepers Review*, *15*, 183–186.
- Raven, J. (1965). *The Coloured Progressive Matrices Test*. London: Lewis Publishers.
- Reitan, R. & Wolfson, D. (1985). *The Halstead-Reitan Neurological Test Battery: Theory and clinical interpretation*. Tucson, AZ: Neuropsychology Press.
- Reynolds, C.R., Chastain, R.L., Kaufman, A.S., & McLean, J.E. (1987). Demographic characteristics and IQ among adults: Analysis of the WAIS-R standardization sample as a function of the stratification variables. *Journal of School Psychology*, *25*, 323–342.
- Royall, D.R., Cabello, M., & Polk, M.J. (1998). Executive dysfunction: An important factor affecting the level of care received by older retirees. *Journal of the American Geriatrics Society*, *46*, 1519–1524.
- Royall, D.R., Mulroy, A.R., Chiodo, L.K., & Polk, M.J. (1999). Clock drawing is sensitive to executive control: A comparison of six methods. *Journal of Gerontology: Psychological Sciences*, *54B*, 328–333.
- Ryan, J.J. & Paolo, A.M. (2001). Exploratory factor analysis of the WAIS-II in a mixed patient sample. *Archives of Clinical Neuropsychology*, *16*, 151–156.
- 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.
- Sobel, M.E. (1988). Direct and indirect effect in linear structural equation models. In J.S. Long (Ed.), *Common problems/proper solutions: Avoiding error in quantitative research* (pp. 46–64). Beverly Hills, CA: Sage.
- Spren, O. & Strauss, E. (1998). *A compendium of neuropsychological tests: Administration, norms, and commentary*. New York: Oxford University Press.
- U.S. Department of Health, Education, and Welfare. (1966). *Equality of educational opportunity*. Washington, DC: U.S. Government Printing Office.
- Wechsler, D. (1997). *Wechsler Adult Intelligence Scale—Third Edition*. San Antonio, TX: Harcourt Brace.
- Welsh, K.A., Fillenbaum, G., Wilkinson, W., Heyman, A., Mohs, R.C., Stern, Y., Harrell, L., Edland, S.D., & Beekly, D. (1995). Neuropsychological test performance in African-American and white patients with Alzheimer's disease. *Neurology*, *45*, 2207–2211.
- Wilkinson, G.S. (1993). *WRAT-3 administration manual*. Wilmington, Delaware: Wide Range.