Education Quality, Reading Recognition, and Racial Differences in the Neuropsychological Outcome from Traumatic Brain Injury

Noah D. Silverberg1,2,*, Robin A. Hanks3, Season C. Tompkins4

1G.F. Strong Rehab Centre, Vancouver, BC, Canada
2Department of Physical Medicine and Rehabilitation, Department of Medicine, University of British Columbia, Vancouver, BC, Canada
3Department of Physical Medicine and Rehabilitation, Wayne State University School of Medicine, Detroit, MI, USA
4Department of Psychology, University of British Columbia, Vancouver, BC, Canada

*Corresponding author at: G.F. Strong Rehab Centre, 4255 Laurel St, Vancouver, BC, Canada V5Z 2G9. Tel.: +1-604-734-1313; fax: +1-604-737-6359.
E-mail address: noah.silverberg@vch.ca (N.D. Silverberg).

Accepted 19 March 2013

Abstract

Ethnically diverse examinees tend to perform lower on neuropsychological tests. The practice of adjusting normative comparisons for the education level and/or race to prevent overpathologizing low scores is problematic. Education quality, as measured by reading recognition, appears to be a more accurate benchmark for premorbid functioning in certain populations. The present study aimed to extend this line of research to traumatic brain injury (TBI). We hypothesized that a measure of reading recognition, the Wechsler Test of Adult Reading (WTAR), would account for racial differences in neuropsychological performance after TBI. Fifty participants (72% African American, 28% Caucasian) with moderate to severe TBI underwent neuropsychological testing at 1-year post-injury. Reading recognition accounted for all the same variance in neuropsychological performance as race and education (together), as well as considerable additional variance. Estimation of premorbid functioning in African Americans with TBI could be refined by considering reading recognition.

Keywords: traumatic brain injury; premorbid functioning; racial differences; neuropsychological testing; psychometrics

Introduction

Compared with Caucasians, African Americans have been found to score lower on a range of neuropsychological tests (e.g., Manly, Jacobs, et al., 1998; Manly, Jacobs, Touradji, Small, & Stern, 2002; Norman, Evans, Miller, & Heaton, 2000; Schwartz et al., 2004; Welsh et al., 1995), raising the risk for false-positive diagnostic errors or overpathologizing in African Americans (Manly et al., 1998; Welsh et al., 1995). Using race-specific normative data in neuropsychological test interpretation has been proposed as an interim solution to this problem (see Manly, 2005; Nabors, Evans, & Strickland, 2000). However, race is a complex phenomenon, not a unitary or biological construct with clear boundaries into which people can be easily grouped (Kaplan & Bennett, 2003; Manly, 2005, 2006; Nabors et al., 2000). In addition, racial differences may have nothing to do with race, as conventionally defined by self-identification. Other confounding variables such as education history may better explain disparities in neuropsychological performance (Manly, 2005, 2006; Nabors et al., 2000; Ryan et al., 2005).

Adjusting normative comparisons for education level (i.e., years of formal education completed) is insufficient to eliminate racial differences, probably because of the historically lower quality of education for many African Americans (Manly, 2005; Nabors et al., 2000; Ryan et al., 2005). In other words, because of disparities in education quality or achievement, years of education may not translate into the same expected ability level for Caucasians and African Americans. Adjusting obtained scores for word recognition performance, considered a proxy measure of education “quality” and its corollaries (e.g., literacy; Manly et al., 2002; Schneider & Lichtenberg, 2011), has emerged as a promising alternative. Manly and colleagues (2002) found that differences between education-matched healthy elderly African Americans and Caucasians on tests of word list learning and memory, figure memory, abstract reasoning, word generation, and visuospatial skills were attenuated by...
controlling for word recognition scores on the Wide Range Achievement Test, Version 3 (WRAT-3). In another study with older non-demented African Americans, WRAT-3 scores strongly predicted neuropsychological performance above and beyond age, sex, years of education, and acculturation (Manly, Byrd, Touradji, & Stern, 2004). Likewise, when elderly Caribbean- and U.S.-born African Americans were compared, WRAT-3 scores were more accurate predictors of neuropsychological test performance than age and years of education (Byrd, Sanchez, & Manly, 2005). In addition, group differences in performance between elderly African Americans and Caucasians on two cancellation tasks were eliminated after equating participants on WRAT-3 scores; the same could not be achieved by matching participants on years of education (Byrd, Touradji, Tang, & Manly, 2004).

These findings in healthy older adults are encouraging but validation in patients with acquired brain dysfunction is necessary before education quality-based normative adjustments can be adopted into clinical practice with ethnically diverse examinees. To date, HIV is the only clinical condition that has been studied. HIV-positive African Americans and Hispanics were more likely than Caucasians to have discrepant education level and quality (indexed by WRAT-3 scores). This discrepancy, but not racial status, helped explain lower neuropsychological performance in ethnically diverse participants (Ryan et al., 2005). Education quality (i.e., WRAT-3 scores) also outperformed education level with regard to the diagnostic accuracy of HIV-associated cognitive impairment in African Americans (Rohit et al., 2007). The explanatory power of education quality in other neurological conditions is not yet known, particularly those with more prominent cognitive sequelae.

The purpose of this study was to examine whether education quality, as indexed by word recognition performance, can account for differences in neuropsychological outcomes from traumatic brain injury (TBI) in adult African Americans and Caucasians. TBI is a leading cause of disability in young- to middle-adulthood (Thornhill et al., 2000; Thurman, Alvers, Dunn, Guerrero, & Sniezek, 1999) and neuropsychological assessment is a key component of TBI clinical management. Minority racial status and lower education level have both been associated with worse outcomes from TBI (Arango-Lasprilla et al., 2008; Gary, Arango-Lasprilla, & Stevens, 2009; Kesler, Adams, Blasey, & Bigler, 2003). The risk of misattributing low neuropsychological test scores to TBI instead of premorbid factors is higher in African Americans due to (on average) poorer education quality. Such overpathologizing can lead to inappropriate clinical decisions (e.g., delayed return to work). Gaining a better understanding of cultural differences and their mediators will help guide clinicians to interpret neuropsychological tests more accurately in ethnically diverse examinees with TBI. Given that word recognition tests are resilient to the effects of TBI and other acquired neurological conditions (Bright, Jaldow, & Kopelman, 2002; Franzen, Burgess, & Smith-Seemiller, 1997; Green et al., 2008; Orme, Johnstone, Hanks, & Novack, 2004; Vanderploeg & Schinka, 2004), they should validly measure the influence of education history even in the presence of severe cognitive impairment. The present study could provide evidence for another application of word recognition tests in TBI, where they already appear useful for estimating premorbid intellectual functioning (Franzen et al., 1997; Green et al., 2008) and predicting post-acute functional abilities, community integration, and level of disability (Hanks, Rapport, Millis, & Deshpande, 1999; Hanks et al., 2008) irrespective of race.

**Method**

Fifty participants in the Southeastern Michigan Traumatic Brain Injury System program were included in this study. All incurred external-force head trauma resulting in (a) post-traumatic confusion duration >24 h, (b) loss of consciousness duration >30 min, (c) Glasgow Coma Scale score <13 in emergency department, and/or (d) trauma-related intracranial neuroimaging abnormalities. They were, on average, 38.1-year old (SD = 14.0) at the time of their injury. More participants were African American (72%) versus Caucasian (28%); 78% were men. Patients in the Traumatic Brain Injury Model System database are described in more detail elsewhere (Corrigan et al., 2007).

As part of a research program, participants underwent neuropsychological testing at 1-year post-injury. Raw scores on six measures, Trail Making Test (Reitan & Wolfson, 1985) Part A and Part B, California Verbal Learning Test, Second Edition (Delis, Kramer, Kaplan, & Ober, 2000) Total Learning and Long Delay Free Recall, phonemic fluency (FAS version) total score, and Grooved Pegboard Test (Reitan & Wolfson, 1985) dominant hand time, were converted to demographically un-adjusted T-scores by (a) looking up the scaled score equivalents in Appendix C of Heaton and colleagues (2004) normative system, normalizing the score distributions and ensuring that higher scores always reflect better performance and then (b) converting the scaled score to a T-score based on the normal distribution to provide a more familiar and precise metric. (Note that the test battery in the Heaton, Miller, Taylor, & Grant, 2004, normative system includes the California Verbal Learning Test (CVLT), not the CVLT-2. However, according to the CVLT-2 technical manual, the two variables analyzed in the present study, Total Learning and Long Delay Free Recall, are equivalent across the CVLT and CVLT-2 versions.) A summary score, the overall test battery mean (OTBM), was then created by averaging these T-scores (sample mean = 38.3 ± 7.1). This approach of creating an OTBM is described elsewhere (Miller & Rohling, 2001). It has been validated in TBI
(Rohling, Meyers, & Millis, 2003) and shown to perform comparably with traditional aggregate indices of neuropsychological impairment (Larrabee, Millis, & Meyers, 2008; Rohling, Williamson, Miller, & Adams, 2003). Participants also completed the Wechsler Test of Adult Reading (WTAR; Wechsler, 2001). Like the WRAT, the WTAR is a word recognition test that requires participants to read a list of words aloud. Although most of the above-reviewed studies used the WRAT, the newer WTAR has the advantages of (a) including only irregularly spelled words, for which correct pronunciation depends on past exposure and learning, possibly making it more resilient to neurological insult (Franzen et al., 1997), and (b) providing predicted WAIS/ [Wechsler Adult Intelligence and Memory Scales- (third edition)(WMS-III)] index scores with confidence intervals, improving linkage to commonly used neuropsychological tests (Wechsler, 2001). The WTAR has been validated in TBI (Green et al., 2008). Scaled scores from the WTAR were analyzed in the present study.

The assessment battery also included the Digit Vigilance Test (Lewis & Rennick, 1979), a measure of attention and processing speed. As well, participants completed self-report measures of psychological distress (Brief Symptom Inventory-18; Derogatis, 2001), problematic alcohol use (Short Michigan Alcoholism Screening Test; Selzer, Vinokur, & van Rooijen, 1982), and substance abuse (Drug Abuse Screening Test; Skinner, 1982).

Results

Demographic and clinical characteristics of the sample are presented in Table 1, as are racial group comparisons on these variables. Caucasian and African American participants had similar demographic profiles with respect to age, education, and gender. The two groups also did not differ on measures of TBI severity, including the Glasgow Coma Scale score at emergency department admission, loss of consciousness duration, and post-traumatic amnesia duration. Finally, Caucasians and African Americans did not differ on measures of psychological distress or problematic alcohol and drug use (Table 1). To confirm racial differences in the neuropsychological outcome after TBI in the present sample, we conducted an analysis of variance with race as the sole fixed factor and the OTBM as the dependent variable. Caucasian participants had higher OTBMs than African Americans, $F(1, 48) = 6.27, p = .016$, Cohen’s $d = 0.84$. This main effect remained significant after controlling for post-traumatic amnesia duration (an indicator of TBI severity), $F(1, 45) = 7.03, p = .011$, Cohen’s $d = 0.85$. As expected, the effect of post-traumatic amnesia severity was also significant in this model, $F(1, 45) = 4.27, p = .045$.

Race (1 = Caucasian, 2 = African American), years of education attainment, and WTAR (scaled scores) were then entered into a linear regression model, with OTBM as the dependent variable. The overall model was significant, $F(3, 43) = 10.26, p < .001, R^2 = .41$. The Bayesian Information Criterion (BIC), an index of model fit that facilitates between-model comparison, was 319.58. Wald’s tests of the beta coefficients demonstrated that only the WTAR contributed unique variance (Table 2). Tolerance values were all >.50 (i.e., acceptable), but an inspection of conditional indices revealed two elevations (>15, the third and fourth dimensions), suggesting possible multicollinearity. The pattern of variance proportions was then inspected;

| Table 1. Demographic and clinical characteristics of Caucasian and African American participants |
|----------------------------------|-----------------|-----------------|-----------------|-----------------|
| Age | Caucasian ($N = 14$) | African American ($N = 36$) | Statistical test value | $p$-value |
| Age | 36.48 | 38.73 | $t = 0.51$ | .620 |
| Education | 14.75 | 13.87 | $t = 1.54$ | .130 |
| Gender | 12.21 | 11.22 | $t = 1.54$ | .130 |
| Time since injury (days) | 2.23 | 1.97 | $t = 1.06$ | .875 |
| Gender | 71% male | 81% male | $\chi^2 = 0.49$ | .484 |
| Time since injury (days) | 3.63 | 3.65 | $t = 0.51$ | .620 |
| Weighted Article Risk Rating | 12.78 | 13.77 | $t = 0.51$ | .620 |
| Loss of consciousness (days) | 3–8 = 43% | 3–8 = 57% | $\chi^2 = 1.02$ | .601 |
| Brief Symptom Inventory (Global Severity Index T score) | 9–12 = 23% | 9–12 = 23% | | |
| Loss of consciousness (days) | 13–15 = 21% | 13–15 = 20% | | |
| Post-traumatic amnesia (days) | 6.58 | 6.31 | $t = 0.07$ | .945 |
| Brief Symptom Inventory (Global Severity Index T score) | 10.19 | 12.17 | $t = 0.93$ | .359 |
| Post-traumatic amnesia (days) | 60.38 | 43.71 | $t = 1.25$ | .219 |
| Brief Symptom Inventory (Global Severity Index T score) | 67.07 | 51.87 | | |
| Post-traumatic amnesia (days) | 59.43 | 54.97 | | |
| Brief Symptom Inventory (Global Severity Index T score) | 12.74 | 10.80 | | |
| Post-traumatic amnesia (days) | 3.08 | 2.09 | Mann–Whitney $U = .759$ | |
| Short Michigan Alcoholism Screening Test | 4.30 | 2.39 | Mann–Whitney $U = .452$ | |
| Drug Abuse Screening Test | 2.15 | 1.08 | Mann–Whitney $U = .452$ | |
| Drug Abuse Screening Test | 3.11 | 1.59 | | |
WTAR and education loaded strongly on the third dimension, whereas WTAR and race loaded on the fourth dimension. Zero-order (non-parametric) correlations confirmed that WTAR was significantly correlated with race, \( r_s(49) = -.44, p = .002 \), and education, \( r_s(49) = .47, p = .001 \), whereas race and education were less correlated with each other, \( r_s(50) = -.26, p = .070 \). This evidence of (mild) multicolinearity suggested that WTAR may be "going after" much of the same variance in the OTBM as race and education. As well, variance in the OTBM in this full (three-predictor) model might have been arbitrarily assigned to WTAR, producing a significant beta weight and misleadingly indicating that it is the most important predictor. WTAR and race/education models were therefore examined separately.

Removing race and education from the full model resulted in non-significant degradation with respect to variance explained, \( \Delta R^2 = -.01, F(2, 45) = 0.53, p = .594 \). Put differently, race and education added essentially no value over and above WTAR in predicting the neuropsychological outcome. The WTAR-only was significant, \( F(1, 47) = 30.35, p < .001, R^2 = .39, BIC = 312.93 \). The magnitude of difference in BIC values between the full and WTAR-only models provided "positive" evidence (by conventional interpretation; Hardin & Hilbe, 2007) that the latter was a better fit to the data. That is, removing race and education resulted in a highly favorable parsimony-for-predictive power tradeoff.

Alternatively, removing WTAR from the full model resulted in a major loss of OTBM variance explained, \( \Delta R^2 = -.18, F(1, 45) = 14.03, p = .001 \). This race and education model was still significant, \( F(2, 46) = 6.53, p = .003, R^2 = .22 \). Its BIC value (336.91) was significantly higher than both the WTAR-only and full models, providing strong evidence that it was the most ill-fitted of the three candidate models.

It is possible that the WTAR explained the same variance in the neuropsychological outcome as race and education plus additional variance, not because it better measures the same underlying construct (education quality), but merely because of shared methods (i.e., the WTAR and OTBM are both performance-based ability tests). To evaluate this alternative explanation, we examined another neuropsychological variable, time to complete the Digit Vigilance Test, which is not part of the OTBM. The Digit Vigilance Test was more modestly correlated with race, \( r_s(49) = -.21, p = .153 \), and education, \( r_s(49) = .30, p = .045 \). It also had minimal multicolinearity with these two variables in a regression model—only race (0.41) and education (0.51) loaded on the sole elevated conditional index (18.96)—while uniquely predicting OTBM (Wald’s test = 3.42, \( p = .001 \)). Moreover, whereas race and education explained no unique variance in the OTBM over and above WTAR scores (as described above), these two variables did explain 11.5% variance over and above Digit Vigilance Test scores, \( F(2, 42) = 4.11, p = .024 \).

**Discussion**

Teasing apart premorbid factors from acquired brain dysfunction is a critical step in neuropsychological test interpretation for patients with TBI. Adjusting normative comparisons for race and/or education level is currently standard practice to avoid misattributing low scores to acquired brain dysfunction in ethnically diverse examiners. However, these demographic variables lack sensitivity to within-race differences and disparities in educational experience. Manly (2005, 2006) and Manly and Echemendia (2007) argue that race should be deconstructed into underlying variables that explain racial differences (e.g., Caucasian > African American) on neuropsychological testing, such as education quality. We found that racial differences in the post-acute neuropsychological outcome from moderate-to-severe TBI were attributable to word recognition performance, a proxy indicator of education quality. Race and education level were essentially unrelated to the TBI outcome, after the influence of word recognition was accounted for. We conclude that education quality may largely explain so-called “racial” differences in the TBI outcome. The findings from this study are consistent with the previous literature demonstrating that correcting for word recognition performance strongly attenuates Caucasian–African American differences in neuropsychological test scores in non-TBI samples, including healthy older adults and patients with HIV (Byrd et al., 2004; Manly et al., 2002; Ryan et al., 2005).

<table>
<thead>
<tr>
<th>Model</th>
<th>Predictors</th>
<th>Unstandardized coefficients</th>
<th>Standard error</th>
<th>Beta coefficients</th>
<th>t</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Race</td>
<td>−1.11</td>
<td>1.97</td>
<td>−0.07</td>
<td>−0.56</td>
<td>.577</td>
</tr>
<tr>
<td></td>
<td>Education</td>
<td>0.38</td>
<td>0.45</td>
<td>0.11</td>
<td>0.85</td>
<td>.399</td>
</tr>
<tr>
<td></td>
<td>WTAR</td>
<td>0.24</td>
<td>0.07</td>
<td>0.54</td>
<td>3.75</td>
<td>.001</td>
</tr>
<tr>
<td>2</td>
<td>WTAR</td>
<td>0.28</td>
<td>0.05</td>
<td>0.63</td>
<td>5.51</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>3</td>
<td>Race</td>
<td>−3.89</td>
<td>2.07</td>
<td>−0.25</td>
<td>−1.88</td>
<td>.066</td>
</tr>
<tr>
<td></td>
<td>Education</td>
<td>1.17</td>
<td>0.45</td>
<td>0.34</td>
<td>2.57</td>
<td>.013</td>
</tr>
</tbody>
</table>

*Note: WTAR = Wechsler Test of Adult Reading.*
As racial differences in neuropsychological performance became apparent, race-stratified normative data emerged as a practical interim solution to counter overpathologizing in African Americans. Our study adds to the mounting literature that education quality is the more relevant construct. Specifically, we demonstrated that this is also true for patients with TBI, using the WTAR scaled score as a measure of word recognition.

Although education quality is difficult to measure directly, especially retrospectively, word recognition tests may be the best available proxy. Considering acculturation (Kennepohl, Shore, Nabors, & Hanks, 2004; Manly et al., 2004; Manly, Miller, et al., 1998) and cultural matches in examinee-examiner dyads (Kennepohl et al., 2004) may further refine neuropsychological assessment of ethnically diverse examinees. Ideally, these variables would have been measured in the present study. Further limiting our findings, we did not screen for symptom invalidity. Although participants were seen in a research setting, it is possible that a non-trivial proportion performed below their capacity, which could have confounded our findings. The effect of education quality may differ across neuropsychological domains (Schneider & Lichtenberg, 2011), but our modest sample size restricted us to an omnibus test of global neuropsychological functioning. Another limitation of our study, and of prior studies using similar methodology, is that word recognition tests may correlate with neuropsychological performance more than education at least in part because of shared method variance (i.e., both involve standardized test-taking). Our supplemental analyses with a neuropsychological predictor that was not part of the OTBM reduced the credibility of this alternative explanation. Race and education lost their predictor power when considered alongside the WTAR, but not when considered alongside an attention task, demonstrating specificity of our main finding.

Given that the results from this study demonstrated that the WTAR accounted for all the same variance in the OTBM as race and education (together), as well as considerable additional variance not explained by these demographic variables, clinicians assembling a neuropsychological test battery for an examinee with TBI may want to include a word recognition test such as the WTAR. This also holds for cognitive research batteries studying outcome after TBI, as word recognition may prove to be a unique factor with regard to functional abilities and community integration after TBI, and may be a more important predictor or covariate than the level of education. Normative data stratified by word recognition performance (transformed to a premorbid intelligence quotient) are becoming increasingly available (e.g., Brooks, Holdnack, & Iverson, 2011), as they have been found to improve diagnostic accuracy even in examinees from the cultural majority (Testa, Winicki, Pearlson, Gordon, & Schretlen, 2009).

The construct(s) measured by word recognition tests such as the WTAR should be clarified in future research. Since decades before their application to understanding racial diversity, such tests have been used to estimate premorbid cognitive ability in neuropsychological assessment (Franzen et al., 1997; Schretlen, Buffington, Meyer, & Pearlson, 2005). Word recognition performance may reflect a combination of innate cognitive ability, cumulative life experience (perhaps not restricted to that gained from formal education), and cultural differences beyond education quality. Moreover, literacy is often used synonymously with education quality in the racial differences literature but may actually be one of several education corollaries.

Funding

This work was supported by the National Institute on Disability and Rehabilitation Research (H133A080044) and Vancouver Coastal Health Research Institute.

Conflict of Interest

None declared.

References


