

Geriatric Performance on an Abbreviated Version of the Boston Naming Test

Angela L. Jefferson

*Alzheimer's Disease Center & Department of Neurology, Boston University
School of Medicine, Boston, Massachusetts, USA*

Sarah Wong and Talia S. Gracer

*Alzheimer's Disease Center & Department of Neurology, Boston University School
of Medicine, Boston, MA and Department of Psychology, Tufts University,
Medford, Massachusetts, USA*

Al Ozonoff

*Department of Biostatistics, Boston University School of Public Health,
Boston, Massachusetts, USA*

Robert C. Green

*Alzheimer's Disease Center & Department of Neurology, Boston University School
of Medicine, Boston, MA and Department of Epidemiology, Boston University
School of Public Health, Boston, Massachusetts, USA*

Robert A. Stern

*Alzheimer's Disease Center & Department of Neurology, Boston University
School of Medicine, Boston, Massachusetts, USA*

Abbreviated neuropsychological protocols are increasingly utilized secondary to time-constraints within research and healthcare settings, yet normative data for these abbreviated instruments are lacking. We present geriatric performances and normative data for the Boston Naming Test 30-item even version (BNT-30). Data were utilized from the BU-ADCC registry (n = 441, ages 55–98) and included 219 normal controls (NC), 155 participants with mild cognitive impairment (MCI), and 67 participants with Alzheimer's disease (AD). The NC group (M = 28.7, SD = 1.8) significantly outperformed both MCI (M = 26.2, SD = 4.4) and AD (M = 22.1, SD = 4.8) groups, and the MCI group outperformed the AD group. Normative data generated for the NC participants revealed a significant between-group difference for sex (males M = 29.1, SD = 1.7; females M = 28.4, SD = 1.8) and race (White M = 28.8, SD = 1.7; African American M = 27.5, SD = 2.1). The racial disparity remained even after adjusting for education level (p = .002) and literacy (p < .001). ANOVAs for the NC group were non-significant for age but significant for education level (p = .001). Geriatric normative data therefore suggest that sex, race, and education are all associated with naming performance, and these variables should be taken into consideration when interpreting geriatric BNT-30 performance.

Key words: Alzheimer's disease, Boston Naming Test, geriatrics, language, lexical retrieval, mild cognitive impairment, neuropsychological measures, normative data

Address correspondence to Angela L. Jefferson, PhD, Alzheimer's Disease Center, Boston University School of Medicine, Robinson Complex, Suite 7800, 715 Albany Street, Boston, MA 02118-2526. E-mail: angelaj@bu.edu

The Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1983) is a widely utilized neuropsychological measure that is sensitive to

detecting compromised lexical retrieval abilities and aphasia through visual confrontation naming. The original 60-item BNT has solid psychometric properties, including strong test-retest reliability (Flanagan & Jackson, 1997) and good concurrent validity (Axelrod, Ricker, & Cherry, 1994). Furthermore, the BNT has been shown to be useful in the differential diagnosis of dementia. For instance, Diehl and colleagues (2005) reported that a combination of Animal Naming and BNT performance correctly distinguished over 90% of patients with frontotemporal dementia and Alzheimer's disease (AD), while a combination of BNT and Mini-Mental State Examination (MMSE) performance correctly distinguished 96% of patients with semantic dementia and AD.

Abbreviated versions of the BNT are potentially useful for several important reasons. Time-constraints in both clinical practice and research settings increasingly favor more efficient neuropsychological instruments. Furthermore, patients with lower levels of education, lower intellectual levels, or severe cognitive impairments are more likely to become frustrated or fatigued during a lengthy protocol (Calero, Arnedo, Navarro, Ruiz-Pedrosa, & Carnero, 2002).

Abbreviated versions of the BNT have comparable psychometric properties when contrasted with the original 60-item version (Fastenau, Denburg, & Mauer, 1998; Williams, Mack, & Henderson, 1989). However, despite increased use, there is limited normative data available for abbreviated versions of the BNT, restricting their clinical and research utility. Mack and colleagues (1992) validated an independently-derived 30-item BNT and confirmed its discriminative utility for dementia, but their normative data were based on a small sample ($n = 26$). Other studies have similarly reported normative data on 30-item versions of the BNT utilizing small sample sizes (Fisher, Tierney, Snow, & Szalai, 1999; Williams et al., 1989).

The present study presents normative data for the BNT 30-item even version (i.e., BNT-30; Williams et al., 1989) according to demographic variables such as age, sex, race, education level, and literacy based on a cohort of 219 geriatric normal controls (NC). We also compare NC performances on the BNT-30 to older adults meeting criteria for two diagnostic categories that are common in the aging population, including mild cognitive impairment (MCI) and AD.

METHODS

Participants

The present study utilized data from the Boston University Alzheimer's Disease Core Center (BU-ADCC) registry, which longitudinally follows older adults with and without memory problems. As part of their annual registry evaluation, participants undergo a comprehensive neurodiagnostic workup, including neurological examination and neuropsychological evaluation. Diagnoses are made at a multidisciplinary consensus consisting of two board certified neurologists, two neuropsychologists (RS, AJ), one nurse practitioner, and other research team members. Inclusion criteria require that participants be age 55 years and older, community-dwelling English speakers, with adequate hearing and visual acuity to participate in the examinations. Exclusion criteria include a history of major psychiatric illness (e.g., schizophrenia, bipolar disorder), neurological illness (e.g., stroke, epilepsy), or head injury with loss of consciousness. The local Institutional Review Board approved data collection efforts for this study, and written informed consent was obtained from all participants prior to testing.

NC participants included 219 elders who, after undergoing the aforementioned neurodiagnostic workup, were designated as cognitively normal. Criteria for inclusion in this group included all objective cognitive performances within the normal range and a Clinical Dementia Rating score (CDR; Morris, 1993) = 0.

MCI participants included 155 individuals meeting widely-accepted research criteria (Petersen, 2004). Inclusion criteria for the *possible* MCI group ($n = 119$) included a decline from previous level of functioning, a lack of dependence in traditionally-defined activities of daily living (Lawton & Brody, 1969), and objective cognitive impairment. The latter criteria was based on neuropsychological data from the BU ADCC registry annual visit with impairment defined as performance falling at least 1.5 SD below available normative data. Inclusion criteria for the *probable* MCI group ($n = 36$) included the same criteria as outlined for possible MCI in conjunction with a subjective (i.e., patient or informant) report of cognitive change(s). Therefore, the distinguishing feature of the probable and possible MCI participants was the presence or absence of a subjective

cognitive complaint, respectively. All MCI participants had a CDR = 0.5.

AD participants included 67 individuals meeting NINCDS-ADRDA criteria (McKhann et al., 1984) for probable ($n = 38$) or possible AD ($n = 34$) with CDR scores ≥ 1.0 .

Neuropsychological Evaluation

Neuropsychological evaluations were conducted by trained psychometricians in a single session. Participants completed a comprehensive protocol encompassing multiple cognitive components, including global cognition, language, verbal and nonverbal visuospatial memory, attention and information processing speed, executive functioning, visuospatial skills, and motor skills. The following selected measures from the protocol are discussed in detail because of their relevance to the current study.

MMSE (Folstein, Folstein, & McHugh, 1975): measures global cognition and ranges 0–30 with lower scores indicating greater general cognitive impairment.

BNT-30 (Williams et al., 1989): measures naming and lexical retrieval abilities. This study utilized an abbreviated 30-item version, including all even items from the original 60-item version (Kaplan et al., 1983). Raw scores range from 0–30, with lower scores indicating greater lexical retrieval difficulties. The total score for this test is the number of correct responses produced spontaneously plus those produced with semantic stimulus cues (Goodglass, Kaplan, & Barresi, 2001).

Wide Range Achievement Test-3 (WRAT-3) Reading Subtest (Wilkinson, 1993): is an achievement measure for reading skills that involves reading aloud words with irregular spelling to sound correspondence (e.g., “benign”). Raw scores are converted to standard scores ranging from 45 to 121 with lower scores reflecting poorer performance. This measure is frequently used to estimate intelligence, and research has demonstrated it is a good measure of literacy and quality of education (Manly et al., 1998).

Geriatric Depression Scale (GDS; Yesavage et al., 1983): assesses depressive symptoms via a 30-item self-report questionnaire. Total scores range from 0 to 30. Scores between 0 and 10 suggest normal mood function and scores

greater than 10 suggest the presence of depressed mood.

Data Analysis Plan

Descriptive statistics and frequencies were generated to summarize demographic (i.e., age, education, sex, and race) and neuropsychological variables (i.e., MMSE, WRAT-3 Reading, GDS). The sample was trichotomized according to diagnostic category (i.e., NC, MCI, AD) and between-group comparisons were made utilizing analyses of variance (ANOVAs) for age, education, GDS, MMSE, and WRAT-3 Reading. Between-group comparisons were made using analysis of covariance (ANCOVAs) for BNT-30 performance, adjusting for relevant variables identified in the prior analysis. Pairwise multiple comparisons were conducted utilizing a Bonferroni adjustment to account for multiple testing.

Among the normal control sample, independent samples t-tests were used to compare BNT-30 performance by sex (i.e., male vs. female), WRAT-3 Reading Subtest (i.e., average vs. high average), and race (i.e., White vs. African American). ANOVAs were employed to evaluate BNT-30 performance among different age groups (i.e., 55–64, 65–74, 75–84, and 85+) and education levels (i.e., high school graduate or less, less than a college graduate, college graduate, less than a graduate degree, graduate degree). Post hoc analyses using Bonferroni correction for multiple comparisons were used to test for significant differences. For all analyses, significance was set a priori at $\alpha = 0.05$.

RESULTS

Demographic Characteristics

Descriptive statistics were calculated for all demographic variables (see Table 1). Participants consisted of 181 males and 260 females with a mean age of 73.2 years ($SD = 8.7$) and mean education of 15.4 years ($SD = 3.1$). The sample was comprised of 81% non-Hispanic White, 18% African American, <1% Native American, and <1% Asian participants.

Between Group Comparisons

Between-group comparisons of the NC, MCI, and AD groups yielded significant differences for

Table 1. Participant Demographic Information and Cognitive Performance

| Variable | NC participants <i>M (SD)</i> <i>n</i> = 219 | MCI participants <i>M (SD)</i> <i>n</i> = 155 | AD participants <i>M (SD)</i> <i>n</i> = 67 |
|-------------------------|--|---|---|
| Age (years) | 71.9 (8.4) | 71.8 (8.0) | 78.2 (6.9) |
| Education level (years) | 16.4 (2.7) | 14.9 (3.2) | 14.0 (2.90) |
| MMSE | 29.2 (1.1) | 28.3 (1.8) | 22.8 (4.4) |
| BNT-30 | 28.7 (1.8) | 26.2 (4.4) | 22.1 (4.8) |
| WRAT-3 reading | 113.2 (6.2) | 105.3 (14.3) | 105.7 (9.0) |
| GDS | 3.8 (4.4) | 4.6 (4.2) | 6.1 (5.9) |
| Sex (% female) | 66 | 55 | 47 |
| Race (% White) | 87 | 71 | 90 |

Note: NC = normal control; MCI = mild cognitive impairment; AD = Alzheimer's disease; *M* = mean; *SD* = standard deviation; MMSE = Mini-Mental State Examination; BNT-30 = Boston Naming Test abbreviated version; WRAT-3 = Wide Range Achievement Test-3; GDS = Geriatric Depression Scale.

age ($F_{(2,425)} = 16.8, p < .0001$) and education level ($F_{(2,425)} = 21.9, p < .0001$). Post-hoc comparisons revealed that the AD group was significantly older than both the NC ($p < .0001$) and MCI ($p < .0001$) groups, and the NC sample was significantly more educated than both the MCI ($p < .0001$) and AD ($p < .0001$) groups.

Significant between-group differences emerged for all neuropsychological variables, including the MMSE, WRAT-3 Reading subtest, and GDS. As expected, there was a significant difference for MMSE score ($F_{(2,425)} = 231.2, p < .0001$) such that the NC group outperformed both MCI ($p < .0001$) and AD ($p < .0001$) groups and the AD group performed significantly worse than the MCI group ($p < .0001$). For WRAT-3 Reading Subtest, there was a significant difference between groups ($F_{(2,425)} = 31.0, p < .0001$) with post hoc comparisons revealing the NC group scored higher than the MCI group ($p < .0001$) and the AD group ($p < .0001$). No significant difference was found between the MCI and AD groups. Adjusting for education, age, WRAT-3 Reading subtest, and GDS, a significant between-group difference was observed for BNT-30 performance ($F_{(2,428)} = 46.8, p < .001$). Post hoc comparisons revealed that the NC group significantly outperformed both MCI ($p = .004$) and AD ($p < .0001$) participants, and the MCI group significantly outperformed AD participants ($p < .0001$), as hypothesized. Finally, there was a significant difference for GDS ($F_{(2,425)} = 6.2, p = .002$) such that the NC group reported significantly less depressed mood than the AD group ($p = .002$); however, there were no significant differences between MCI participants and NC or AD participants. It is noteworthy that the mean GDS score was well within normal limits for all three groups (see Table 1).

BNT-30 Normative Data by Sex, Race, Age, and Education

Among the cognitively normal controls ($n = 219$), BNT-30 normative data revealed that males outperformed females ($t_{(217)} = 2.5, p = .01$; see Table 2), and non-Hispanic Whites outperformed African-Americans ($t_{(216)} = 3.8, p = .0002$; see Table 2). Follow-up analyses were conducted to elucidate the disparities. When education level was included as a covariate, the sex differences were no longer significant ($F_{(1,218)} = 3.0, p = .09$). In contrast, the racial disparity remained ($F_{(1,215)} = 6.1, p = .01$), even after adjusting for several possible confounding variables including education level ($F_{(1,215)} = 5.6, p = .02$), a proxy of education level (i.e., WRAT-3 Reading; $F_{(1,215)} = 8.1, p = .005$), sex ($F_{(1,215)} = 4.7, p = .03$), and age ($F_{(1,215)} = 6.0, p = .02$).

No significant difference was noted for age ($F_{(3,219)} = 2.3, ns$; see Table 2). However, a significant difference was found for education level ($F_{(4,219)} = 7.3, p < .0001$; see Table 2). Post hoc comparisons found that high school educated participants performed significantly worse than those with a college degree ($p = .001$), some graduate schooling ($p < .0001$), and a graduate degree ($p < .0001$).

A significant difference was found for WRAT-3 Reading level ($F_{(2,216)} = 11.0, p = .00003$), with those participants achieving a high average reading score performing significantly better on the BNT-30 than those achieving an average score ($p = .0006$).

Recommendations for BNT-30 Normative Data Cut-Off Scores

For clinical and research recommendations, BNT-30 cut-off scores were calculated based

BNT-30 NORMATIVE DATA

Table 2. *BNT-30 Normative Data by Sex, Race, Age, or Education*

| Sex | <i>N</i> | Min | Max | <i>M (SD)</i> | Cut-off (–1.5 <i>SD</i>) | Cut-off (–2.0 <i>SD</i>) |
|-----------------------|----------|-----|-----|---------------|---------------------------|---------------------------|
| Male | 74 | 22 | 30 | 29.1 (1.7) | 27 | 26 |
| Female | 145 | 22 | 30 | 28.4 (1.8) | 26 | 25 |
| Race ^a | | | | | | |
| White | 191 | 22 | 30 | 28.8 (1.7) | 26 | 25 |
| African American | 27 | 22 | 30 | 27.5 (2.1) | 24 | 23 |
| Age | | | | | | |
| 55–64 years old | 47 | 24 | 30 | 29.0 (1.3) | 27 | 26 |
| 65–74 years old | 83 | 22 | 30 | 28.8 (1.8) | 26 | 25 |
| 75–84 years old | 73 | 22 | 30 | 28.2 (2.0) | 25 | 24 |
| 85+ years old | 16 | 26 | 30 | 28.9 (1.4) | 27 | 26 |
| Education level | | | | | | |
| High school or less | 29 | 22 | 30 | 27.3 (2.3) | 24 | 23 |
| Some college | 40 | 22 | 30 | 28.3 (1.9) | 25 | 25 |
| College grad | 40 | 25 | 30 | 28.9 (1.5) | 27 | 26 |
| Less than grad degree | 63 | 23 | 30 | 29.0 (1.6) | 27 | 26 |
| Graduate degree | 47 | 25 | 30 | 29.1 (1.2) | 27 | 27 |

Note: ^a*n* = 218, one participant was categorized as “Other” and omitted from the racial between group analyses; cut-off scores are intended to reflect values where at least mild (–1.5 *SD*) or moderate (–2.0 *SD*) impairment is present in reference to our normative sample.

Table 3. *BNT-30 Normative Data by Age and Race**

| Age range | <i>N</i> | Min | Max | <i>M (SD)</i> | Cut-off (–1.5 <i>SD</i>) | Cut-off (–2.0 <i>SD</i>) |
|------------------|----------|-----|-----|---------------|---------------------------|---------------------------|
| White | | | | | | |
| 55–64 years old | 38 | 26 | 30 | 29.2 (1.1) | 28 | 27 |
| 65–74 years old | 73 | 24 | 30 | 29.0 (1.5) | 27 | 26 |
| 75–84 years old | 66 | 22 | 30 | 28.4 (2.0) | 25 | 24 |
| 85+ years old | 14 | 26 | 30 | 28.8 (1.5) | 27 | 26 |
| African American | | | | | | |
| 55–64 years old | 9 | 24 | 30 | 28.1 (2.0) | 25 | 24 |
| 65–74 years old | 10 | 22 | 30 | 27.0 (2.6) | 23 | 22 |
| 75–84 years old | 7 | 26 | 29 | 27.0 (1.0) | 26 | 25 |
| 85+ years old | 1 | 30 | 30 | 30.0 (0.0) | – | – |

Note: **n* = 218, one participant was categorized as “Other” and omitted from the racial between group analyses; cut-off scores are intended to reflect values where at least mild (–1.5 *SD*) or moderate (–2.0 *SD*) impairment is present in reference to our normative sample.

Table 4. *BNT-30 Normative Data by Age and WRAT-3 Reading**

| Age range | <i>N</i> | Min | Max | <i>M (SD)</i> | Cut-off (–1.5 <i>SD</i>) | Cut-off (–2.0 <i>SD</i>) |
|------------------------|----------|-----|-----|---------------|---------------------------|---------------------------|
| Average | | | | | | |
| 55–64 years old | 21 | 26 | 30 | 28.8 (1.3) | 27 | 26 |
| 65–74 years old | 15 | 23 | 30 | 27.4 (2.1) | 24 | 23 |
| 75–84 years old | 11 | 22 | 29 | 26.7 (2.1) | 26 | 25 |
| 85+ years old | – | – | – | – | – | – |
| High average or better | | | | | | |
| 55–64 years old | 25 | 26 | 30 | 29.3 (0.9) | 28 | 28 |
| 65–74 years old | 67 | 22 | 30 | 29.1 (1.6) | 27 | 26 |
| 75–84 years old | 61 | 22 | 30 | 28.5 (1.8) | 26 | 25 |
| 85+ years old | 15 | 26 | 30 | 29.0 (1.4) | 27 | 26 |

Note: **n* = 215, 4 participants did not complete the WRAT-3; cut-off scores are intended to reflect values where at least mild (–1.5 *SD*) or moderate (–2.0 *SD*) impairment is present in reference to our normative sample.

Table 5. *BNT-30 Normative Data by Age and Sex*

| Age range | <i>N</i> | Min | Max | <i>M (SD)</i> | Cut-off ($-1.5 SD$) | Cut-off ($-2.0 SD$) |
|-----------------|----------|-----|-----|---------------|-----------------------|-----------------------|
| Male | | | | | | |
| 55–64 years old | 13 | 24 | 30 | 29.1 (1.7) | 27 | 26 |
| 65–74 years old | 29 | 23 | 30 | 29.4 (1.4) | 27 | 27 |
| 75–84 years old | 24 | 22 | 30 | 28.8 (2.1) | 26 | 25 |
| 85+ years old | 8 | 27 | 30 | 28.6 (1.4) | 27 | 26 |
| Female | | | | | | |
| 55–64 years old | 34 | 26 | 30 | 28.9 (1.2) | 27 | 27 |
| 65–74 years old | 54 | 22 | 30 | 28.5 (1.9) | 26 | 25 |
| 75–84 years old | 49 | 22 | 30 | 27.9 (1.9) | 25 | 24 |
| 85+ years old | 8 | 26 | 30 | 29.1 (1.5) | 27 | 26 |

Note: Cut-off scores are intended to reflect values where at least mild ($-1.5 SD$) or moderate ($-2.0 SD$) impairment is present in reference to our normative sample.

on $-1.5 SD$ and $-2.0 SD$ to reflect at least mild or moderate impairment, respectively. The rationale for presenting multiple cut-off scores is related to the empirical emphasis of defining cognitive impairment as $-1.5 SD$ below peers for characterization of MCI. These cut-off scores are organized according to sex (Table 2), race (Table 2), age (Table 2), education level (Table 2), age and race (Table 3), age and WRAT-3 Reading score (Table 4), age and sex (Table 5), as well as age and education (Table 6). Because the demographic

composition of each clinician's patient base may vary widely, we provide numerous normative data tables that allow clinicians to choose the appropriate combination of demographic variables that best meets individual client-base needs.

DISCUSSION

Among cognitively normal elders, our data suggest sex, race, and education level are all associated

Table 6. *BNT-30 Normative Data by Age and Education*

| Education | <i>N</i> | Min | Max | <i>M (SD)</i> | Cut-off ($-1.5 SD$) | Cut-off ($-2.0 SD$) |
|---------------------------|----------|-----|-----|---------------|-----------------------|-----------------------|
| 55–64 years old | | | | | | |
| High school or less | 3 | 26 | 28 | 27.3 (1.2) | 26 | 25 |
| Some college | 7 | 28 | 30 | 29.3 (0.8) | 28 | 28 |
| College grad | 10 | 28 | 30 | 29.6 (0.7) | 29 | 28 |
| Less than graduate degree | 16 | 24 | 30 | 28.8 (1.6) | 26 | 26 |
| Graduate degree | 11 | 26 | 30 | 28.9 (1.4) | 27 | 26 |
| 65–74 years old | | | | | | |
| High school or less | 10 | 22 | 30 | 27.1 (3.1) | 22 | 21 |
| Some college | 13 | 24 | 30 | 28.3 (1.7) | 26 | 25 |
| College grad | 14 | 25 | 30 | 28.5 (1.8) | 26 | 25 |
| Less than graduate degree | 24 | 26 | 30 | 29.5 (0.9) | 28 | 28 |
| Graduate degree | 22 | 25 | 30 | 29.3 (1.2) | 28 | 27 |
| 75–84 years old | | | | | | |
| High school or less | 12 | 22 | 29 | 27.0 (1.9) | 24 | 23 |
| Some college | 17 | 22 | 30 | 27.9 (2.2) | 25 | 24 |
| College grad | 13 | 25 | 30 | 28.6 (1.6) | 26 | 25 |
| Less than graduate degree | 22 | 23 | 30 | 28.6 (2.1) | 25 | 24 |
| Graduate degree | 9 | 28 | 30 | 29.0 (0.9) | 28 | 27 |
| 85+ years old | | | | | | |
| High school or less | 4 | 27 | 30 | 28.5 (1.7) | 26 | 25 |
| Some college | 3 | 26 | 30 | 28.0 (2.0) | 25 | 24 |
| College grad | 3 | 29 | 30 | 29.7 (0.6) | 29 | 29 |
| Less than graduate degree | 1 | 30 | 30 | 30 (0.0) | – | – |
| Graduate degree | 5 | 27 | 30 | 29.0 (1.2) | 27 | 27 |

Note: Cut-off scores are intended to reflect values where at least mild ($-1.5 SD$) or moderate ($-2.0 SD$) impairment is present in reference to our normative sample.

with performance on an abbreviated version of the BNT. However, the sex differences appear to be secondary to educational disparities among males and females among our older cohort. This difference is not completely surprising, and it likely reflects a generational effect with older males obtaining higher levels of education than older females. Previous research has noted a similar disparity on the 60-item BNT, such that males outperform females in both healthy (Ross & Lichtenberg, 1998; Welch, Doineau, Johnson, & King, 1996) and AD samples (Ripich, Petrill, Whitehouse, & Ziol, 1995). Previous studies have attributed such sex disparities to differing neural networks mediating language among males and females, as females have greater bihemispheric representation for language than males (Vikingstad, George, Johnson, & Cao, 2000). Superficially, our data corroborate the previous studies, as the elderly NC males outperformed females. However, after further analyses of the data, we found that education may explain the sex differences. Future studies assessing naming performances may wish to include education as a covariate when examining sex differences.

The racial-group differences for BNT-30 performance noted in our study are consistent with some earlier studies, as White elders reportedly obtain higher abbreviated BNT scores than African Americans (Manly et al., 1998; Manly, Jacobs, Touradji, Small, & Stern, 2002). These racial disparities have, in part, been explained by educational achievement or literacy differences. For instance, racial discrepancies noted on comprehensive neuropsychological protocols are generally eliminated when educational achievement (Manly et al., 1998) or a proxy measure of educational quality/literacy (i.e., WRAT-3 Reading subtest) is considered (Manly et al., 2002). However, it is important to note that previous work has shown that not all racial group differences are attenuated when a proxy measure is considered (Manly et al., 2002). In contrast to most prior literature, the racial discrepancies noted in the current study remained after the WRAT-3 Reading subtest was included as a covariate. A plausible explanation for the racial disparity is that our chosen proxy measure does not fully encompass the complex set of variables (e.g., socioeconomic factors, cultural experience) that impact educational quality or literacy in this cohort (Manly et al., 2002). Therefore, researchers and clinicians should exercise caution

in relying solely on literacy measures to adjust for racial disparities, as reading level may not fully account for racial differences on some neuropsychological measures.

Older individuals with at least some college education outperformed individuals with a high school education or less. These data are not surprising, as the extant literature contains multiple examples of similar associations between education and performance on the original 60-item BNT (Fox, Warrington, Seiffer, Agnew, & Rossor, 1998; Welch et al., 1996). Previous studies have suggested that this education difference may be due to an increase in performance variability among individuals with less than a high school education (Welch et al., 1996). Our data support this variability theory, as those individuals with less than a high school education had a larger standard deviation for BNT-30 performance as compared to the more educated participants.

The lack of association between age and BNT-30 performance contradicts previous research utilizing the 60-item BNT in a geriatric cohort (Ross & Lichtenberg, 1998). The discrepancy between our findings and previous research may be secondary to differences in sample demographics, including racial composition and education achievement. More specifically, 81% of our sample is White compared to 44% from a prior study by Ross and Lichtenberg (1998). The mean education achievement of our sample was approximately four years greater than that of Ross and Lichtenberg (1998). Another explanation for the differences may be that our normative sample was carefully examined to exclude persons with MCI or early symptoms of dementia, which are more common in older samples. Additional research is warranted to clarify an association, if any, between age and naming performance.

The findings from this study augment the extant literature in several ways. First, we present robust normative data for the BNT 30-item even version based on a large sample of healthy controls. Previous studies reporting normative data for abbreviated BNT versions have utilized much smaller sample sizes (Fisher et al., 1999; Mack et al., 1992; Williams et al., 1989), but our normative sample consisted of more than 200 participants. Using this larger sample size, we were able to provide breakdowns according to various demographic variables, including age, education level, education quality/literacy, and race. Furthermore,

participants within our sample underwent comprehensive neurodiagnostic work-ups to confirm their normal control status, which increases the likelihood that our normal controls are cognitively and functionally normal.

Despite the numerous strengths of the present study, two limitations must be considered. First, the majority of our sample is comprised of non-Hispanic White individuals (81%); therefore, clinicians using the education or sex breakdown to interpret BNT-30 performance for other racial groups should exercise caution, as our findings suggest there is some racial disparity in BNT-30 performance. Future research should focus on presenting normative data stratified across even larger sample sizes for racial minorities and identifying specific cultural and linguistic variables that may affect performance. Second, our sample underwent a thorough neurodiagnostic evaluation to ensure that participants were normal controls. Therefore, the normative data presented in this study may reflect a “super” geriatric sample rather than something observed in an epidemiological study, which limits the generalizability of our findings.

In summary, the present study compared performances on the BNT 30-item even version among NC, MCI, and AD participants and presented geriatric normative data. Our findings suggest that sex, race, education level, and education quality/literacy are associated with BNT performance; therefore, when interpreting naming performance on this abbreviated measure, normative data or statistical adjustment for these factors should be considered.

ACKNOWLEDGMENTS

This research was supported by F32-AG022773 (ALJ), K12-HD043444 (ALJ), R03-AG026610 (ALJ), R03-AG027480 (ALJ), R01-AG009029 (RCG), K24-AG027841 (RCG), R01-HG/AG002213 (RCG), P30-AG013846 (Boston University Alzheimer’s Disease Core Center), and M01-RR00533 (General Clinical Research Centers Program of the National Center for Research Resources, NIH).

REFERENCES

- Axelrod, B., Ricker, J., & Cherry, S. (1994). Concurrent validity of the MAE visual naming test. *Archives of Clinical Neuropsychology*, *9*(4), 317–321.
- Calero, M. D., Arnedo, M. L., Navarro, E., Ruiz-Pedrosa, M., & Carnero, C. (2002). Usefulness of a 15-item version of the Boston Naming Test in neuropsychological assessment of low-educational elders with dementia. *J Gerontol B Psychol Sci Soc Sci*, *57*(2), P187–191.
- Diehl, J., Monsch, A. U., Aebi, C., Wagenpfeil, S., Krapp, S., Grimmer, T., et al. (2005). Frontotemporal dementia, semantic dementia, and Alzheimer’s disease: The contribution of standard neuropsychological tests to differential diagnosis. *J Geriatr Psychiatry Neurol*, *18*(1), 39–44.
- Fastenau, P. S., Denburg, N. L., & Mauer, B. A. (1998). Parallel short forms for the Boston Naming Test: Psychometric properties and norms for older adults. *Journal of Clinical and Experimental Neuropsychology*, *20*(6), 828–834.
- Fisher, N. J., Tierney, M. C., Snow, W. G., & Szalai, J. P. (1999). Odd/Even short forms of the Boston Naming Test: Preliminary geriatric norms. *Clin Neuropsychol*, *13*(3), 359–364.
- Flanagan, J. L., & Jackson, S. T. (1997). Test-retest reliability of three aphasia tests: Performance of non-brain-damaged older adults. *J Commun Disord*, *30*(1), 33–42; quiz 42–43.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”. A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198.
- Fox, N. C., Warrington, E. K., Seiffer, A. L., Agnew, S. K., & Rossor, M. N. (1998). Presymptomatic cognitive deficits in individuals at risk of familial Alzheimer’s disease. A longitudinal prospective study. *Brain*, *121*(Pt 9), 1631–1639.
- Goodglass, H., Kaplan, E., & Barresi, B. (2001). The nature of the deficits. In *The assessment of aphasia and related disorders* (3rd ed., pp. 5–11). Philadelphia: Lippincott Williams & Wilkins.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *The Boston naming test* (2nd ed.). Philadelphia: Lea & Febiger.
- Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: Self-maintaining and instrumental activities of daily living. *The Gerontologist*, *9*(3), 179–186.
- Mack, W. J., Freed, D. M., Williams, B. W., & Henderson, V. W. (1992). Boston Naming Test: Shortened versions for use in Alzheimer’s disease. *J Gerontol*, *47*(3), 154–158.
- Manly, J. J., Jacobs, D. M., Sano, M., Bell, K., Merchant, C. A., Small, S. A., et al. (1998). Cognitive test performance among nondemented elderly African Americans and whites. *Neurology*, *50*(5), 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*(3), 341–348.
- McKhann, G., Drachman, D., Folstein, M., Katzman, R., Price, D., & Stadlan, E. M. (1984). Clinical diagnosis of Alzheimer’s disease: Report of the NINCDS-ADRDA Work Group under the auspices of Department of Health and Human Services Task Force on Alzheimer’s Disease. *Neurology*, *34*(7), 939–944.
- Morris, J. C. (1993). The Clinical Dementia Rating (CDR): Current version and scoring rules. *Neurology*, *43*(11), 2412–2414.

BNT-30 NORMATIVE DATA

- Petersen, R. (2004). Mild cognitive impairment as a diagnostic entity. *J Intern Med*, 256(3), 183–194.
- Ripich, D. N., Petrill, S. A., Whitehouse, P. J., & Ziol, E. W. (1995). Gender differences in language of AD patients: A longitudinal study. *Neurology*, 45(2), 299–302.
- Ross, T., & Lichtenberg, P. (1998). Expanded normative data for the Boston Naming Test for use with urban, elderly medical patients. *The Clinical Neuropsychologist*, 12(4), 475–481.
- Vikingstad, E. M., George, K. P., Johnson, A. F., & Cao, Y. (2000). Cortical language lateralization in right handed normal subjects using functional magnetic resonance imaging. *Journal of the Neurological Sciences*, 175(1), 17–27.
- Welch, L., Doineau, D., Johnson, S., & King, D. (1996). Educational and gender normative data for the Boston Naming Test in a group of older adults. *Brain and Language*, 53(2), 260–266.
- Wilkinson, G. S. (1993). *Wide range achievement test-3 (WRAT-3) administration manual*. San Antonio, TX: The Psychological Corporation.
- Williams, B. W., Mack, W., & Henderson, V. W. (1989). Boston naming test in Alzheimer's disease. *Neuropsychologia*, 27(8), 1073–1079.
- Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., et al. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research*, 17(1), 37–49.

Copyright of Applied Neuropsychology is the property of Lawrence Erlbaum Associates and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.