Clinical Aphasiology, Vol. 24, 1996, pp. 135–144

The Influence of Demographic Variables on Normal Elderly Subjects' Performance on the Boston Naming Test

Michael L. Kimbarow, Stephen J. Vangel, Jr., and Peter A. Lichtenberg

The Boston Naming Test (BNT; Kaplan, Goodglass, and Weintraub, 1983), an established clinical tool for assessing naming deficits associated with a variety of neuropathological conditions, is a widely used test of confrontation naming abilities of aphasic adults. Thus, knowledge of age-related performance on naming tasks for non–brain-damaged elderly individuals is necessary in order to distinguish between normal aging and neurological dysfunction.

Studies using the Boston Naming Test to explore the effects of aging on naming abilities have generally found decreases in these abilities particularly beyond the eighth decade of life. In a widely referenced study, VanGorp, Satz, Kiersch, and Henry (1986) established BNT norms for adults ranging in age from 59 to 95. Using a sample of 78 patients, VanGorp et al. reported a modest decline in naming ability with age ($\underline{r} = -.33$). However, the authors noted that their group was selected from a "super normal group of older adults" (p. 704) and that their results were uncontaminated by sociocultural factors that might lower overall scores. Mitrushina and Satz (1989) also reported a modest but significant relationship between advancing age and performance on the BNT ($\underline{r} = -.29$, $p < .001$). VanGorp et al. and Mitrushina and Satz reported that despite significant correlations with age, mean scores varied by only five points between those aged 57 to 65 and those over 80, although the variance did increase with age. Thus normative data presented suggested that cutoff scores change little with advancing age.

In 1989, Nicholas, Brookshire, MacLennan, Schumacher, and Porrazzo published revised administration and scoring procedures for the BNT and provided norms obtained from 60 non–brain-damaged adults. In assessing their subjects' performance, Nicholas et al. found no relationship between age and total BNT score ($r = 0.10$, $p > .05$) generally consistent with the findings of VanGorp et al. (1986) and Mitrushina and Satz (1989). Nicholas et al. also found a weak but significant relationship ($r = 0.30$, $p < 0.01$) between education and total BNT score.

Remarkably, the mean BNT scores reported in these three studies were almost identical. On close examination this is not surprising as the mean education level of the subjects used in all three studies was also nearly identical (approximately two years of college). All three studies appeared to draw their subjects from a highly educated cohort.

We note that none of these previous investigations provided data on the ethnic mix of their subjects. This leads to a logical question. Are previously published BNT norms valid for a culturally diverse, less-educated population? Clinical experience with patients served at our hospital suggest they are not. Similar to those at many inner city health care facilities, our patient population is multiethnic, of low socioeconomic status (SES), and is generally less educated than their middle SES counterpart. We have observed that many of our patients routinely miss many BNT items and that these errors appear to be unrelated to any neurological deficit. Consequently, we believe it would be clinically useful to explore factors unrelated to neurological status which may influence performance on this test.

One such investigation was reported by Brookshire and Nicholas (1995). They analyzed the relationship between aphasic adults' performance on individual BNT items and their respective U-values (estimated frequency of a word in print per million tokens; Carroll, Davies, and Richman, 1971). They reported that an item's frequency of occurrence in printed materials was not strongly related to their aphasic subjects' ability to name that item. To our knowledge there has been no effort to explore whether BNT item U-values influence naming performance of normal adults.

To understand some of the non-neurological factors which may influence performance on the BNT, and to generate normative data appropriate to our clinical population, the present investigation was designed to address the following questions:

1. Is there a relationship between age, race, and education in cognitively intact elderly adults' performance on the Boston Naming Test?
2. Is there a relationship between word frequency in printed material (U-value) for Boston Naming Test items and demographic factors in performance by cognitively intact elderly adults on the Boston Naming Test?

METHOD

Subjects

Subjects were 113 consecutive admissions between 1991 and 1993 to a geriatric rehabilitation unit in a Midwestern university hospital. The sample consisted of 67 African-Americans ranging in age between 60 and 90 ($\underline{M}$ = 74.77, $\underline{SD}$ = 5.98) and ranging between 3 and 18 years of education ($\underline{M}$ = 10.44, $\underline{SD}$ = 3.44). There were 46 Caucasian subjects ranging from 63 to 91 years of age ($\underline{M}$ = 75.33, $\underline{SD}$ = 6.60) and ranging between 8 and 18 years of education ($\underline{M}$ = 11.02, $\underline{SD}$ = 2.03).

To rule out cognitive dysfunction all subjects were administered the Dementia Rating Scale (Mattis 1988) and scored at or above the cutoff score of 123 based on Mattis' 1988 norms. Subjects received the Wechsler Memory Scale-R Logical Memory Subtest and all scores were within one standard deviation of the mean based on Ivnik et al.'s (1992) norms. None of the subjects scored above the cutoff of 10 on the Geriatric Depression Scale (Yesavage et al. 1983). Approximately one-third of the subjects had hip fractures secondary to a fall, one-third had knee replacements secondary to arthritis, and one-third were admitted for reconditioning following lengthy illness. Postsurgical orthopedic patients were tested 2 to 3 weeks after surgery and were on no narcotic medications when tested. All subjects were independent in Activities of Daily Living (ADL) upon discharge from the hospital.

Procedure

Subjects were individually administered the Boston Naming Test as part of a larger neuropsychological battery given by a licensed neuropsychologist with more than 10 years of experience. The test was administered and scored according to the standardized procedures indicated in the BNT Scoring Booklet. Correct responses were credited only if the patient spontaneously provided the picture name listed in the BNT Scoring Booklet within 20 seconds or provided the picture name within 20 seconds following the listed stimulus cue. Responses were scored on-line by the examiner and recorded in the test booklet.

Scoring Reliability

The only correct responses accepted by the examiner were those listed in the BNT Scoring Booklet; thus, there were no instances of disagreement on the scoring of individual items between the examiner and first author.

To determine the reliability of the overall BNT score assigned to each patient, the senior author reviewed the test protocols of all 113 patients and independently computed the total BNT score for each patient. Point-to-point percentage reliability for the overall score was calculated using the formula:

$$\frac{\text{Number of agreements}}{\text{Number of agreements and disagreements}} \times 100$$

There was 100% agreement between the examiner's scores (PL) and the author's (MK) scores for each patient.

RESULTS

The mean number of items correct for our 113 subjects ranged from 13 to 59 years of age ($\underline{M}$ = 40.55, $\underline{SD}$ = 11.74). To explore the relationship between age, education, race, and total BNT score, we computed Pearson and point bi-serial correlations between demographic characteristics and total BNT score (Table 1). Results showed that all three demographic variables were significantly correlated to BNT score.

Stepwise multiple regression of BNT scores on education, race, and age showed that all three variables significantly predicted BNT score ($\underline{F}$ (3,113) = 15.26, $\underline{p} < .001$). Education alone accounted for 16% of the BNT variance; education and race accounted for 25% of the variance; and all three demographic variables predicted 29% of the variance in BNT scores. With a normative population, test theory predicts total error variance. Our findings suggest we can predict almost a third of the BNT variance when we factor in our three demographic variables. Further, stable predictor beta weights (education = .34, race = .29, age = −.26) were found for each demographic

Table 1. Pearson and Point Bi-Serial Correlations between Demographic Characteristics and Boston Naming Test Scores

	Age	*Race*	*Education*	*BNT*
Age	—	.04	−.06	−.22*
Race		—	.14	.35*
Education			—	.40*

*$\underline{p} < .05$
Note: The Boston Naming Test (Kaplan et al. 1983)

variable entered. Thus, it is critical with the BNT to factor in demographic variables when interpreting scores. Consequently, normative data on the BNT, dividing groups into age, race, and education cohorts are provided in Table 2. Cutoff scores for normal performance are defined as two standard deviations below the mean.

Three two-way Analyses of Variance (ANOVAs) were done to examine the effects of race (African-American and Caucasion) and education (≤ 12 yrs. vs. > 12 yrs.) on performance of BNT item subsets trichotomized into groups based on the highest 20 U-values, middle 20 U-values, and lowest 20 U-values. Age accounted for 4% of the variance and was not included as a variable in our analyses. Alpha was adjusted to .01 for each analysis in order to control for Type I error. Results for the 20 BNT items with the highest U-values showed scores were significantly different by race ($F(1,109) = 8.35$, $p < .01$), but not by education group level ($F(1,109) = 2.40$, $p > .05$) or by race and education interaction ($F(1,109) = .456$, $p > .05$). Results for the 20 BNT items with midrange U-values were also significantly different by race ($F(1,109) = 15.04$, $p < .001$), but not by education group ($F(1,109) = 3.51$, $p > .05$) or their interaction ($F(1,109) = 1.48$, $p > .05$). Results for the 20 BNT items with lowest U-values also were significantly different by race ($F(1,109) = 15.63$, $p < .001$), with a trend observed for education level ($F(1,109) = 4.88$, $> p > .01$). The interaction between race and education was not significant ($F(1,109) = 2.00$, $p > .05$). In all three analyses, African-American subjects scored lower than Caucasian subjects.

DISCUSSION

The results of this study suggest that use of previously published Boston Naming Test (BNT) norms obtained from highly educated normal adults is not valid when evaluating performance of less-educated urban dwellers. Education alone accounted for 16% of BNT scores variance in our stepwise multiple regression.

Our findings are generally in agreement with previous studies of the BNT which show a modest decline in naming ability associated with advancing age. This decline may be due to a reduction in speed of response (Albert, Heller, and Milberg, 1988) or perhaps to declines in lexical retrieval (Nicholas, Obler, Albert, and Goodglass, 1985).

We note that age along with education and race added 4% to the total BNT Score variance. Thus, age alone is not a strong explanatory variable to account for the performance of our subjects.

To our knowledge our results are the first indication that older adults' performance on the BNT may also be influenced by one's racial and educational

Table 2. Means, Standard Deviations, and Ranges for Boston Naming Test Scores for African-American and Caucasian Cognitively-Intact Adults across Two Age Groups and Two Educational Cohorts

		African-American		*Caucasian*	
Education		*Age*		*Age*	
		60–74 ***(N=17)***	***75+*** ***(N=18)***	***60–74*** ***(N=12)***	***75+*** ***(N=11)***
≥12	Mean	35.13	30.44	43.75	40.45
	SD	11.18	9.99	10.28	11.80
	Range	14–50	15–47	24–56	21–56
		(N=19)	***(N=13)***	***(N=12)***	***(N=11)***
≥12	Mean	45.0	37.08	48.92	48.45
	SD	12.76	3.59	6.40	6.01
	Range	14–59	13–50	40–56	38–57

background. We observed that African-American subjects scored lower than Caucasian subjects when BNT items were divided into groups based on U-values, or estimated frequency of a word in print per million tokens (Carroll et al., 1971). The difference between African-Americans and Caucasians did not result only from the obvious moderator variable, years of education. Other variables that may explain our finding include socioeconomic status, and/or ethnicity factors. Reliable information on these variables is not always easily accessible. However, race and cultural factors, while not explanatory variables, need to be considered by clinicians when interpreting BNT scores of their adult patients.

Table 3 provides an item-by-item comparison of naming accuracy on the BNT as reported by Nicholas et al. (1989) and the present investigation. Nicholas et al.'s findings are based on modified administration and scoring procedures. They observed that it would be valid to use their norms when using standard administration procedures as long as "the responses specified in the BNT Scoring Booklet are accepted as correct." (p. 577).

Using the Nicholas et al. standard, we identified BNT items missed by 10% or more of subjects in both studies. In the present study, three of the first 20 BNT test items fall below the 10% cutoff compared to one in the Nicholas et al. investigation. Close examination of our subjects' performance suggests that cultural/educational effects are most likely captured in the middle 20 items of the test. Starting with BNT item # 22 (snail), 19 of middle 20 BNT items were missed by 10% or more of our subjects as compared to four items missed by 10% or more of the subjects in the Nicholas et al. study. The final 10 BNT items were almost equally difficult for subjects in both investigations.

If there were no cultural/educational effects we would expect African-American subjects and Caucasian subjects to respond identically throughout the entire test irrespective of U-value status. There were differences between the highly educated subjects used in the Nicholas et al. investigation and the subjects used in the present study. This finding supports the use of our revised norms when testing subjects drawn from a similar inner-city cohort.

The need for accurate norms to use when evaluating BNT performance of neurologically impaired adults is clear. The results of the present investigation provide data on BNT performance by a group of less-educated, cognitively intact elderly adults living in an urban environment. Consequently, we believe these norms should be considered when evaluating BNT performance of aphasic patients drawn from a similar inner-city population. In support of this position, we observe that if we were to use Nicholas et al. (1989) suggested BNT cutoff score of 48 with the subjects in our investigation, 82% of them would be identified as impaired, contrary to their neuropsychological test profiles. Future studies should compare and contrast BNT scores of older, racially diverse, cognitively-intact subjects, and demented subjects.

Table 3. A Comparison of Percent Correct Naming Accuracy for BNT Items Reported by Nicholas et al. (1989) and Kimbarow, Vangel, and Lichtenberg (1996)

BNT Item	*Nicholas et al. (1989)* *% Correct*	*Kimbarow et al. (1996)* *% Correct*
1. Bed	100	100
2. Tree	100	100
3. Pencil	100	99
4. House	90*	100
5. Whistle	100	97
6. Scissors	98	100
7. Comb	100	100
8. Flower	98	98
9. Saw	100	100
10. Toothbrush	98	99
11. Helicopter	97	91
12. Broom	100	100
13. Octopus	100	85*
14. Mushroom	93	91
15. Hanger	100	96
16. Wheelchair	100	100
17. Camel	98	96
18. Mask	100	84*
19. Pretzel	100	78*
20. Bench	100	96
21. Racquet	100	92
22. Snail	100	78*
23. Volcano	95	73*
24. Seahorse	92	52*
25. Dart	98	74*
26. Canoe	98	75*
27. Globe	92	70*
28. Wreath	100	88*
29. Beaver	82*	42*
30. Harmonica	87*	83*
31. Rhinoceros	83*	64*
32. Acorn	93	73*
33. Igloo	98	54*
34. Stilts	100	75*
35. Dominoes	93	80*
36. Cactus	100	65*
37. Escalator	100	74*
38. Harp	100	82*
39. Hammock	98	57*
40. Knocker	90*	66*
41. Pelican	90*	44*

(Continued)

Table 3. A Comparison of Percent Correct Naming Accuracy for BNT Items Reported by Nicholas et al. (1989) and Kimbarow, Vangel, and Lichtenberg (1996) (continued)

BNT Item	*Nicholas et al. (1989)* *% Correct*	*Kimbarow et al. (1996)* *% Correct*
42. Stethoscope	97	50*
43. Pyramid	95	51*
44. Muzzle	88*	53*
45. Unicorn	88*	22*
46. Funnel	95	65*
47. Accordion	92	65*
48. Noose	87*	50*
49. Asparagus	93	66*
50. Compass	47*	22*
51. Latch	57*	43*
52. Tripod	82*	34*
53. Scroll	90*	42*
54. Tongs	82*	46*
55. Sphinx	80*	15*
56. Yoke	73*	27*
57. Trellis	88*	40*
58. Palette	70*	25*
59. Protractor	35*	04*
60. Abacus	55*	10*

Note: Items marked with * denote BNT items missed by 10% or more of the subjects in either study.

In conclusion, we agree with the observation of VanGorp et al. (1986), that sociocultural factors may contaminate and lower performance on the BNT. We also agree with the suggestion of Nicholas et al. (1989), that "when brain-damaged adults misname items on the BNT one should be cautious in attributing these errors to the patient's brain damage." (p. 577). We would add that when brain-damaged adults from culturally and educationally diverse milieus misname or fail to name BNT items, one needs to consider fully the patients' racial and educational background before attributing poor performance to neurological dysfunction.

REFERENCES

Albert, M. S., Heller, H. S., & Milberg, W. (1988). Changes in naming abilities with age. *Psychology and Aging, 3*, 173–178.

Brookshire, R. H., & Nicholas, L. E. (1995). Relationship of word frequency in printed material and judgement of word frequency in daily life to Boston

Naming Test performance of aphasic adults. In M. Lemme (Ed.), *Clinical aphasiology* (Vol. 23, pp. 107–120). Austin, TX: PRO-ED.

Carroll, J. B., Davies, P., & Richman, B. (1971). *The American Heritage Word Frequency Book.* New York: Houghton Miflin.

Ivnik, R. J., Malec, J. F., Smith, E. G., Tangalos, R. C. Peterson, R. C., Kokmen, E., & Kurland, L. T.(1992). Mayo's older American normative studies: WMS–R norms for ages 56 to 94. *The Clinical Neuropsychologist, 6,* supplement, 49–52.

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

Mattis, S. (1988). *Dementia Rating Scale: Professional Manual.* Odessa, FL: Psychological Assessment Resources.

Mitrushina, M., and Satz, P. (1989). Differential decline of specific memory components in normal aging. *Brain Dysfunction, 2,* 330–335.

Nicholas, M., Obler, L., Albert, M. & Goodglass, H. (1985) Lexical retrieval in healthy aging. *Cortex, 21,* 595–606.

Nicholas, L. E., Brookshire, R. H., MacLennan, D. L., Schumacher, J. G., & Porrazzo, S. A. (1989). Revised administration and scoring procedures for the Boston Naming Test and norms for non–brain-damaged adults. *Aphasiology, 6,* 569–580.

VanGorp, W., Satz, P., Kiersch, M., & Henry, R. (1986). Normative data on the Boston Naming Test for a group of normal older adults. *Journal of Clinical and Experimental Neuropsychology, 8,* 702–705.

Yesavage, J. A., Brink, T. L., Rose, T. L., Lum, O., Huang, V., Adey, M., & Leirer, O. (1983). Development and validation of a geriatric depression screening scale: A preliminary report. *Journal of Psychiatric Research, 17,* 37–49.