Boston Naming Test: Normative data for older Australians

Olivier Piguet

Centre for Education and Research on Ageing at Concord Hospital and Prince of Wales Medical Research Institute, Australia

Joanne L. Millar

Centre for Education and Research on Ageing at Concord Hospital, Australia

Hayley P. Bennett

Centre for Education and Research on Ageing at Concord Hospital and Prince of Wales Medical Research Institute, Australia

Tanya C. Lye, Helen Creasey

Centre for Education and Research on Ageing at Concord Hospital, Australia

G. Anthony Broe

Centre for Education and Research on Ageing at Concord Hospital and Prince of Wales Medical Research Institute, Australia

Individuals aged over 80 years represent the fastest growing segment of the population. It is becoming increasingly important to investigate the effect of age on cognitive functions such as language, in order to document "normal" and "abnormal" functioning. A task commonly used to test naming ability in clinical practice is the Boston Naming Test (BNT). Although norms exist for this age group, they may have limited applications because of small sample sizes on which they were derived. In addition, this test uses stimulus items that have been shown to be culturally specific. This study presents normative data for the BNT for two levels of education and two age bands based on a randomly selected Australian sample of older adults between the age of 81 and 94 years. Frequencies of the most common error types made in this group of non-demented individuals are also reported.

The proportion of the Australian population aged over 85 years is predicted to rise from 1.3% in 1999 to over 5% in 2051 (Australian Bureau of Statistics, 1999). Also, there is an exponential increase in age-related prevalence of dementia, with rates doubling every 5 years in individuals over 65 years of age (Jorm, Korten & Henderson, 1987) reaching 10% above 75 and 30–40% above 85 years in a Sydney-based study (Waite, Broe, Creasey, et al., 1995). As a corollary, an increasing number of individuals over the age of 80 years will at some stage undergo neuropsychological evaluation with many found to exhibit some degree of cognitive impairment. Even in the absence of neurodegenerative disorders of ageing, many cognitive functions, including language abilities, decline subtly but significantly in later life. A common and frustrating complaint of older adults is word-finding difficulty (LaBarge, Edwards & Knesevich, 1986). Along with memory

Address for correspondence: Olivier Piguet, Centre for Education and Research on Ageing, Concord Hospital C25, Concord, NSW 2139, Australia. E-mail: olivierp@psych.usyd.edu.au

impairment, naming deficits are considered to be one of the first signs of the very early stages of Alzheimer's Disease (AD; Jacobs, Sano, Dooneief, et al., 1995; Morris, McKeel, Storandt, et al., 1991; Storandt & Hill, 1989) and a decline in naming has been found to be highly sensitive to disease progression (Knesevich, LaBarge & Edwards, 1986).

A common task to examine naming ability in normal ageing and dementia is the Boston Naming Test (BNT; Kaplan, Goodglass & Weintraub, 1983). Naming ability as measured by this test has been demonstrated to remain relatively stable with advancing age in healthy older adults (LaBarge et al., 1986). Cross-sectional studies using English speakers report a significant difference in performance between groups but not before 70 or 75 years of age (Albert, Heller & Milberg, 1988; Borod, Goodglass & Kaplan, 1980; Howieson, Holm, Kaye et al., 1993; Nicholas, Brookshire, MacLennan, et al., 1985; Welch, Doineau, Johnson & King, 1996) accompanied by an increased inter-individual variability in naming performance (Neils, Baris, Carter et al., 1995; VanGorp, Satz, Kiersch & Henry, 1986). Longitudinal data also indicate a similar decline greatest for subjects over the age of 70 years (Au, Joung, Nicholas & Obler,1995). This decline has been found to be independent of measures of visual impairment (Howieson et al., 1993) and has also been reported in French- (Colombo-Thuillard & Assal, 1992), and Dutch-speaking (Mariën, Mampaey, Vervaet, et al., 1998) but not Spanish-speaking subjects (Allegri, Mangone, Villavicencio, et al., 1997).

Different sets of BNT norms are available for elderly subjects or patients. However, existing normative studies suffer from some important limitations. First, normative data for the group of 80+ years are based on very small samples ranging between 5 and 52 subjects (LaBarge et al., 1986; Neils et al., 1995). A notable exception comes from the Mayo group (Ivnik, Malec, Smith et al., 1996) who published norms on American individuals aged 55 years and over using a sample that included 126 subjects over the age of 80 years. Second, highly educated subjects tend to be over-represented. For example, Van Gorp et al.'s (1986) 80+ years group had a mean education level of 15.2 years. With few exceptions (e.g. Lichtenberg, Ross & Christensen, 1994), educational level has been found to significantly correlate with BNT performance, particularly in the oldest population (Welch et al., 1996) whether cognitively intact or not (Henderson, Frank, Pigatt et al., 1998; Neils et al, 1995; Ross & Lichtenberg, 1997; Tombaugh and Hubley, 1997). Also, reported age by education interactions suggest that participants with education beyond high school level show less decline in naming ability with age compared to those with less than high school education (Neils et al., 1995; Welch et al., 1996). Thus, norms uncorrected for education may overestimate the extent of the naming deficit in participants with limited formal education.

Third, strict medical exclusion criteria used in most studies have resulted in the selection of optimally healthy individuals or "super normals" (Albert et al., 1988; Howieson et al., 1993; LaBarge et al., 1986) who may not be representative of the general population of the age group. Ross and colleagues (Lichtenberg et al., 1994; Ross, Lichtenberg & Christensen, 1995) provide normative data on a demographically diverse North American sample of neurologically intact geriatric inpatients with a number of medical ailments (e.g. hypertension, diabetes, hypothyroidism and orthopaedic injuries) in an attempt to overcome the limitation of other studies to generalise to "urban medical populations". The mean BNT scores reported in their samples were lower than other studies suggesting that medical illness impacted upon cognitive functioning, in the absence of a clinically identifiable neurodegenerative process. Finally, most available BNT norms were derived from North American samples each with different demographic characteristics and thus might not be applicable to other countries. An Australian study by Worrall, Yiu, Hickson and Barnett (1995) produced norms which fell 2–5 points below the mean scores for North American adults of similar age (Farmer, 1990; Nicholas et al., 1989; Van Gorp et al., 1986). Unfortunately, their study included only 20 subjects over 80 years. It is unclear whether this difference was due to demographic characteristics (e.g. education), or to cultural differences in the difficulty of some of the BNT items administered. It cannot be assumed that Australian elderly perform at a similar level on the BNT as their North American counterparts.

In addition to providing quantitative information on BNT performance, many studies have investigated the qualitative nature of naming ability and naming deficits in both ageing and dementia, analysing the types of errors most commonly made in these groups. These data have been used for normative purposes (reporting the most common error types found in "normal" older adults), and to provide information about the possible cognitive mechanisms which break down as naming deficits arise or progressively worsen. Several distinct cognitive processes are required for successful naming. These include: (a) visual perception of the target; (b) activation of the semantic concepts associated with the object; (c) identification of, and access to, the label corresponding to the semantic concepts within the lexical network; and finally (d), articulation of the name via activation of motor speech programs (Crowe, Dingjan & Helme, 1997; Nicholas, Obler, Albert & Goodglass, 1985). As a breakdown at any level of these processes may result in naming impairment, error types have been analysed in terms of their visual, semantic or lexical/phonemic nature. Visual errors occur at an initial processing stage and are typically misperceptions of an object (e.g. "fountain pen" for *asparagus*). Semantic errors are varied, and include subordinate, within-category or superordinate errors, or vague associations related to the target. All these errors indicate a difficulty accessing the correct conceptual information from the semantic system to provide an accurate response (e.g. "lettuce" "vegetable" or "you eat it" for *asparagus*). Circumlocutory errors and detailed descriptions of the target indicate that the necessary semantic information about the target is available but that access to the lexical network to generate the corresponding label is somehow blocked. Phonemic errors may represent partial retrieval of a label from the lexical network leading to impaired articulation of the object name (e.g. "apagus" for *asparagus*).

Semantic errors such as circumlocutions and semantically-related associations seem to increase in frequency with age compared to visuo-perceptual errors or phonemic errors (Albert et al., 1988; Chenery, Murdock & Ingram, 1996). It is commonly accepted that elderly subjects have access to semantic knowledge about objects they identify but have difficulty accessing the lexical network which provides the object's name (Bowles & Poon, 1985; Nicholas et al., 1985). The prevalence of visual errors in normal ageing, and whether they are the result of peripheral visual impairment or perceptual breakdown remains somewhat uncertain. Nicholas et al. (1985) reported no differences in visual errors with age whereas Albert et al. (1988) reported a significant increase in visual errors after the age of 70 years, and concluded that naming difficulties may be multiply determined in older people. In a longitudinal study, Au et al. (1995) also implicated multiple processes (lexical, semantic and perceptual processes) in the changes in naming ability across the life span. In an Australian study of healthy individuals ranging between the age of 55 and 92 years (mean age: 70.5 years), semantically related errors were most frequently reported (36%), followed by "don't know" responses (35%), descriptions (i.e. circumlocutions; 12%) and misperceptions (10%). Semantic associative errors, unrelated errors and phonemic errors together accounted for less than 8% of errors (Worral et al., 1995). This study did not report whether the nature of the naming errors changed with advancing age.

In summary, many studies have investigated naming ability in both healthy and demented elderly individuals but normative data on the BNT are minimal for persons in their ninth and tenth decades of life. The relative lack of adequate norms for this age segment probably arises because it currently comprises only a small percentage of the population and with many individuals being medically unwell, unable or unwilling to be tested. However, the importance for normative studies on this section of the community cannot be underestimated given its rapid projected growth over the next decades and the need for identification of early language change signalling possible incipient dementing processes. Furthermore, there is a need for norms that are culturally specific, comprising participants who are non-demented but who may be exhibiting aged-related illnesses such as hypertension, heart disease, peripheral vascular disease and high cholesterol that have been found to be common in the oldest-old (e.g. Waite, Broe, Creasey et al., 1997). Normative data based on such a sample will be more representative of the general population and have a greater applicability to the individuals referred to clinical/hospital settings for cognitive assessment, as they will be demographically similar.

This study aims to provide normative data based on a sample of randomly selected non-demented Australians aged 80 years and over who present with a variety of medical illnesses common to their age group. These norms will provide scaled scores and percentile equivalents of raw scores on the BNT broken down for two levels of education and for two age groups. A secondary aim is to investigate common naming errors made by this sample. This analysis will offer valuable information regarding the nature of the naming deficit in this age segment of the Australian population.

Method

Participants

Participants were selected from the Sydney Older Persons Study (SOPS). SOPS is a longitudinal study following 630 randomly selected community dwellers from the inner west of Sydney, Australia, who were initially aged 75 years and over at the time of recruitment in 1991 and reviewed every 3 years. This study was designed to investigate various aspects of ageing and dementia including: the impact of systemic and neurodegenerative illnesses; brain correlates of ageing; risk factors associated with "successful" versus "unsuccessful" ageing; psychosocial quality of life; and other outcome data (Creasey, Waite, Grayson, et al., 2001). There were 299 remaining participants at the time of the 6-year review. For the present study, only those who obtained a score of 0 ("well") or 0.5 ("questionable dementia") on the Clinical Dementia Rating (CDR; Hughes, Berg, Danziger, et al., 1982) and for whom education information was available were included. Participants were not excluded if they presented with medical illnesses common to this age group such as hypertension, heart disease or peripheral vascular disease nor were they excluded on the basis of poor vision (as peripheral visual impairment is relatively common in this age group). However, acuity was measured using a Snellen chart and presence/absence of visual impairment recorded for each participant. Forty-four women and 51 men (mean age = 85.2 years; range = 81–94.5 years) were included in this study.

Procedure

The BNT was part of a large test battery covering all major cognitive domains that was administered over two sessions. Because of constraints related to the testing procedures used at the time of the baseline and the 3-year review assessments, the BNT was not administered in full in one session. Thirty-two items were administered on one occasion, the remaining 28 items being administered during the second session. The mean time interval between the two sessions was six weeks (*SD* = 5 weeks). The order of testing remained constant for all participants. Administration of the test deviated slightly from the standardised instructions in that all 60 items were administered starting with item 1. However, the BNT scores were computed according to the standardised rules of administration for starting and stopping points, thus making the scores compatible with other published normative results. Another change from the original administration procedure was that, if the item could not be named, a letter cue was provided (e.g. "it starts with the letter M" for *mask*) before the phonemic cue was given. The provision of letter cues before the phonemic cue did not affect the final score, as it was calculated according to the scoring instructions (i.e. the number of spontaneously correct items plus the number of correct items following a semantic cue).

Any erroneous responses from Item 1 through 60 on the BNT were recorded and coded according to the detailed classification system of Hodges, Salmon and Butters (1991). Only the first erroneous response to the target item (i.e. before semantic, letter or phonemic cueing) was coded and included in the error analysis. This is the same coding procedure used by Hodges et al. (1991), and others (e.g. LaBarge et al., 1992).

The CDR (Hughes et al., 1982) is a short questionnaire designed to rate changes in functional abilities in six behavioural and cognitive domains including memory, orientation, judgment and problem solving, involvement in community affairs, involvement in home and hobbies, and personal care. These ratings serve to determine the presence and severity of dementia in an individual. In this study, the CDR was administered to an informant nominated by each participant. The informant CDR has been found to sensitively discriminate between "well" and "demented" individuals (Waite et al., 1998). In addition, others have demonstrated that informants generally appraise and report accurately cognitive and functional deficits in dementia patients, even those with very mild or "questionable" dementia (Koss, Patterson, Ownby et al., 1993; Morris et al., 1991).

Results

The sample demographic characteristics, including Mini Mental State Examination score (Folstein, Folstein & McHugh, 1975), and descriptive statistics for the BNT are presented in Table 1. Four participants obtained a score of less than 32 points (30, 29, 28 and 22 points), which corresponds to a performance of at least 1.5 standard deviation below the means based on the Mayo norms for this age group (Ivnik et al., 1996). The effects of age, education and gender on BNT score were examined using correlational and step-wise multiple regression analyses to determine their respective contribution to the overall BNT score. Modest but significant correlations were found between BNT score and level of education ($r = .28$; $p < .01$), and between BNT score and age ($r = -.23$; $p < .05$).

Frequencies for the different types of errors are shown in Figure 1. Visual errors were the

TABLE 1
Descriptive Statistics for Sociodemographic and Boston Naming Test Variables ($N = 95$)

Variable	Mean	*SD*	Range
Age	85.4	2.9	81–94
Years of education	10.37	2.06	6–19
Mini Mental State Examination score	27.0	2.3	19–30
Boston Naming Test total	48.6	8.5	22–60
Boston Naming Test errors	13.8	8.1	0–38

most common (25.8%), followed by ambiguous visual/semantic (19.6%), semantic associative (13.4%), within-category (12.2%) and circumlocutory errors (11.4%). Non-responses constituted 11.5% of overall errors. The other error types: semantic superordinate, phonemic, perseverative and unrelated errors were very infrequent accounting altogether for less than 6% of the errors. The t tests revealed a significant effect of visual acuity as measured by the Snellen chart on visual errors ($t = -2.55$, $df = 93$; $p < .05$) but not on the other types. Errors were not uniformly distributed across the items. For example, *pretzel* accounted for almost a quarter of all visual errors (often mistaken for a snake or a knot). Similarly, 24% of the ambiguous visual/semantic errors were made on *beaver* (often mistaken for a rat, platypus or wombat).

For the norming exercise, participants were *first* divided into two groups according to the number of years of education, namely less than 10 years ($n = 40$) and 10 or more years ($n = 55$). In a second instance, participants were divided into two age groups: less than 86 years ($n = 57$) and 86 years or older ($n = 38$). These groupings were based on the correlational analyses, on the group average years of education (about 10 years) and age (approximately 85 years) and an attempt to maintain an adequate number of subjects in each group. Subdivisions into smaller age brackets were not attempted due to the relatively narrow age range of the sample and the desire to keep cell sizes to a maximum. As expected, t tests revealed a significant difference between the two education groups for BNT score ($t = -3.96$, $df = 93$; $p = .02$) but not for age, visual acuity or CDR rating. Similarly, there was a significant difference in BNT scores between the two age groups ($t = -2.13$, $df = 93$; $p = .03$) but not for education,

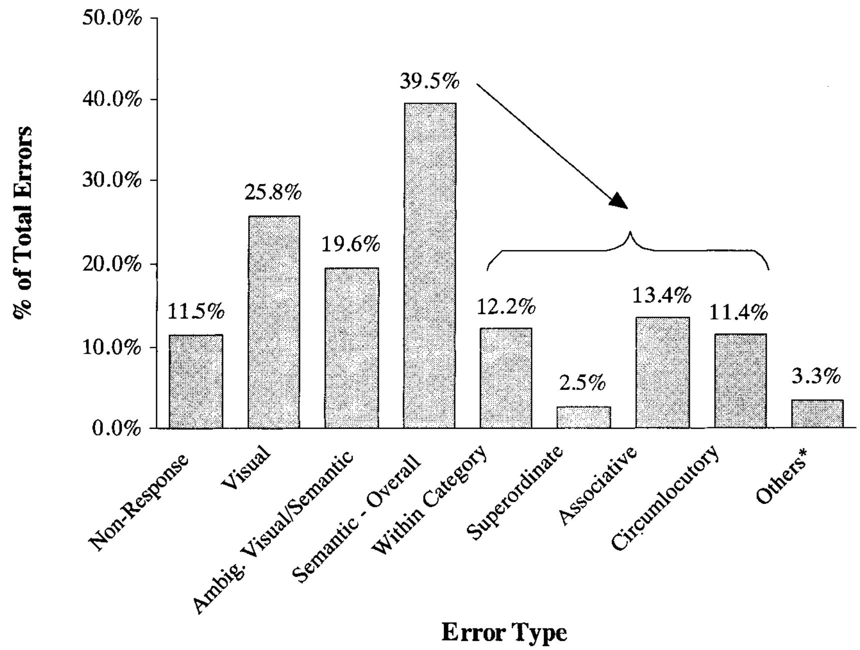

FIGURE 1
Proportion of each BNT error type in a sample of non-demented individuals aged 81–94 years ($N = 95$).
*Others include phonemic errors (0.90%), perseverations (0.01%) and unrelated responses (2.4%).

TABLE 2
Percentiles and Scaled Score Equivalents for BNT Raw Scores With Adjustment for Education (Age range: 81–94 years)

Percentiles	BNT Raw scores* < 10 years (*N* = 40)	BNT Raw scores* ≥ 10 years (*N* = 55)	Scaled scores
<1		*< 28*	2
1	*< 22*	28	3
2	22	29	4
3–5	*23–29,* 30	*30–31*,32	5
6–10	*31–34,* 35–37	34–39	6
11–18	38–39	*40,* 41–45	7
19–28	40–42	46–48	8
29–40	43–44	49–51	9
41–59	45–48	52–53	10
60–71	49–50	54–55	11
72–81	51–53	56	12
82–89	*54, 55–56*	*57*	13
90–94	*57*	58	14
95–97	*58, 59*	59	15
98	*60*	60	16
99			17
>99			18

Note:* Numbers in italics represent smoothing of raw data.

TABLE 3
Percentiles and Scaled Score Equivalents for BNT Raw Scores With Adjustment for Two Age Groups (Age range: 81–94 years).

Percentiles	BNT Raw scores* < 86 years (*N* = 57)	BNT Raw scores* ≥ 86 years (*N* = 38)	Scaled scores
<1	*<29*		2
1	29	*<22*	3
2	*30–31*, 32	22	4
3–5	*33–37*, 38	*23–27*, 28	5
6–10	39–40	*29,* 30–34	6
11–18	41–43	35–37	7
19–28	44–46	38–40	8
29–40	*47*, 48–49	*41*, 42–45	9
41–59	50–53	*46*, 47–50	10
60–71	54–55	*51*, 52	11
72–81	56	53–56	12
82–89	57–58		13
90–94	59	57	14
95–97		58	15
98	60	59	16
99		*>59*	17
>99			18

Note:* Numbers in italics represent smoothing of raw data.

visual acuity or CDR rating. Raw scores for each group were first converted to a percentile rank (based upon the cumulative frequency distribution) to ensure that scaled scores were normally distributed. Each percentile rank was then converted to its corresponding z-score from the standard normal distribution, and then to a scaled score (with a mean of 10 and a standard deviation of 3). Education-, and age-adjusted percentiles and scaled scores are presented in Table 2 and Table 3 respectively. The smoothing of data reflects the interpolation of scores which were not obtained in this sample.

Discussion

This study aimed to provide culture-, as well as age-specific norms for the BNT, a test widely used to detect language impairment in clinical practice. Norms do exist for elderly people but are generally based on very small, North American samples and may therefore not necessarily offer reliable information. The sample size in this study is between twice and five times larger than most other BNT normative studies for this age group. Unlike other normative studies which have applied very strict inclusion and exclusion criteria, this study presents norms based on a randomly selected sample of elderly community dwellers over the age of 81 years who may have been experiencing medical conditions commonly associated with ageing. Consequently, these norms offer a greater relevance to clinicians in their daily practice. Furthermore, the provision of additional norms is an attempt to address the issue of cultural specificity between Northern American populations and other English-speaking countries creating differences in the level of difficulty for some BNT items. This is an important concern as the use of inappropriate normative data to evaluate performance on this test may result in an overestimate of language impairment.

Worrall and colleagues (1995) have examined the issue of cultural difference in detail and reported a mean BNT score in their Australian sample that was between 2 and 5 points lower than scores for similar age groups from American norms, the oldest groups (80–84 and 85+ years) showing the largest difference. Using the standardised 60-item version of the BNT, the present study reveals a similar finding with an overall mean BNT score approximately 1 point higher than those reported by Worrall et al. (1995) for their 20 participants aged 80–84 and 85+. Furthermore, when a cut-off score of two standard deviations below the mean percentage correct for all items is applied, the same 3 items identified by Worrall et al. (*pretzel, beaver* and *protractor*) are again found to be the least often named correctly. Whilst *protractor* is the second last item in the list and is therefore among the most difficult, the common failure to name *pretzel* and *beaver,* items 19 and 29 respectively, suggests other causes most likely cultural.

Interestingly, their attempt at correcting this cultural difference by offering alternative items (such as *platypus* and *pizza* in replacement of *beaver* and *pretzel* respectively) resulted in no significant improvement in total BNT score although the mean success on these items was closer to their American counterparts and similar to neighbouring items. Similarly, we examined the correlation between the total BNT score in this sample for all items from 1 through 60 and the score obtained through standardised administration in an attempt to determine the impact of these unusual items. The obtained correlation was very high (r = .98) with an average score difference of 0.94 points (range of –9 to 6; the minus sign indicates a standard score *less than* the true score) and only 8 participants exhibiting a difference in excess of 3 points. The positive difference indicates that on average, the score based on standardised stopping and starting points yielded a slightly higher number of correct items than the total score. This finding would support the recommended administration procedures despite the presence of the unusual stimuli. Whether the range of possible score differences is important enough to warrant a full presentation will obviously depend upon the clinical situation and question under investigation. However, this high correlation also suggests that cultural differences between Australian and North American norms are not explained solely on the basis of the three unusual items. This finding notwithstanding, the existing difference indicates that clinicians should exert caution in administering abbreviated forms of the BNT where the items have been selected based on their increasing difficulty according to the original order of the BNT (Mack, Freed, Williams & Henderson, 1992). The use of such short forms may inflate the difference between obtained and real scores. Certainly, these findings underscore the importance of norms that are culturally specific and appropriate to the sample, or patient groups, under consideration.

These data allow the provision of normative tables for two levels of education as well as for two age bands. Admittedly, a weakness of these norms is that no concurrent correction for age and education is offered. This subdivision was not attempted in an effort to maintain the maximum cell size. Although some accuracy may have been lost in the process, this has been traded for greater reliability. A second possible weakness of these results is associated with the task presentation. In this study, all 60 BNT items were administered over two sessions and this differs from the standard presentation procedure. It is possible that performance on this task depends in part upon progressive familiarity with the task and the items and this may have been lost with the two separate sessions. However, this appears to be unlikely for two reasons. First, within each session, the presentation was ordered from the easiest to the hardest items and progressive familiarity was therefore always possible. Second, participants to this study were all non-demented individuals. Recollection of the task, if not of the items, and therefore familiarity to the demands associated with the task between the two sessions was possible.

The variety of existing systems of error analysis prevents a direct comparison but the pattern of errors obtained in this study is generally similar to that found elsewhere (Bayles, Tomoeda & Trosset, 1990; Goodglass, 1980). Rare in younger age groups, perceptual errors are not infrequent in this age group. It is uncertain whether these reflect peripheral deficits or sensory disturbance (such as impaired vision due to cataracts or macular degeneration), or a disruption of central processes. A considerable contribution to this type of errors arises from the misperception of the culturally-biased item *pretzel*. However, it is difficult to attribute the higher rate of visual errors solely to a deficit in perceptual processing of a single item. Similarly, *beaver* has the largest single contribution to the ambiguous visual/semantic error category. These items, which may have skewed the error classification, outline again an inherent difficulty in interpreting this test in an Australian population, and could arguably be excluded from the error analysis. Thus, the two highest error categories are inflated by items that are very unfamiliar to this Australian sample.

Also commonly observed are "don't know" responses. In AD patients, these responses are thought to represent a loss of semantic information to the point that sufficient semantic information is not available to produce even a related response (LaBarge et al., 1992). However, Chertkow and Bub (1990) found that for about 5-10% of "don't know" responses in their AD sample, probe questions indicated that semantic memory for the concept was relatively intact. In this study, no specific instructions were given to the participants to try and attempt to name each drawing or to use "don't know" only if they genuinely had no idea about a particular stimulus. These errors are therefore difficult to interpret as participants may have responded "don't know" when in fact they may have been aware of the general class of the stimulus (e.g. the animal class for *beaver*) but failed to verbalise it. For each item, only the first error was scored and an initial "don't know" may have been subsequently changed to a different, possibly correct, answer following the presentation of the semantic cue. This suggests that some of the "don't know" responses may not be linked with core semantic breakdown in some participants. Finally, of note is that the BNT contains a high proportion of items not easily classifiable into standard superordinate categories (such as *globe*, *muzzle*, *unicorn*, *scroll*, *wreath* or *stilts*) and this may account for the relative lack of superordinate responses (Hodges et al., 1991).

Summary

A number of issues regarding the use of the BNT in elderly populations have been addressed and investigated within this study. Firstly, normative data for two education groups and two age bands have been provided based on a large group of randomly selected Australians, filling a gap in the current BNT literature. In addition to these norms, frequencies of error types are reported describing the most common errors made in this group of non-demented individuals.

Acknowledgements

This research was supported in part by a grant from the National Health and Medical Research Council of Australia and an Infrastructure Stream C grant from the Department of Health of New South Wales, Australia. We would like to thank Jessica Carroll and Enid Sawley for their participation in data collection.

References

Albert, M.S., Heller, H.S., & Milberg, W. (1988). Changes in naming ability with age. *Psychology and Aging, 3(2),* 173–178.

Allegri, R.F., Mangone, C.A., Villavicencio, A.F., Rymberg, S., Taragano, F.E., & Baumann, D. (1997). Spanish Boston Naming Test norms. *The Clinical Neuropsychologist, 11(4),* 416–420.

Au, R., Joung, P., Nicholas, M., & Obler, L.K. (1995). Naming ability across the adult life span. *Aging and Cognition, 2(4),* 300–311.

Australian Bureau of Statistics (1999). *Population projections 1999 to 2101.* Canberra: ABS (Cat. No. 3220.0)

Bayles, K.A., Tomoeda, C.K., & Trosset, M.W. (1990). Naming and categorical knowledge in Alzheimer's Disease: The process of semantic memory deterioration. *Brain and Language,* 39, 498–510.

Borod, J.C., Goodglass, H., & Kaplan, E. (1980). Normative data on the Boston Diagnostic Aphasia Examination, Parietal Lobe Battery, and the Boston Naming Test. *Journal of Clinical Neuropsychology, 2(3),* 209–215.

Bowles, N.L., & Poon, L.W. (1985). Aging and retrieval of words in semantic memory. *Journal of Gerontology, 40(1),* 71–77.

Chenery, H.J., Murdock, B.E., & Ingram, J.C.L. (1996). An investigation of confrontation naming performance in Alzheimer's dementia as a function of disease severity. *Aphasiology, 10(5),* 423–441.

Chertkow, H., & Bub, D. (1990). Semantic memory loss in dementia of Alzheimer's type. *Brain, 113,* 397–417.

Colombo-Thuillard, F., & Assal, G. (1992). Adaptation française du test de dénomination de Boston Versions abrégées. *European Review of Applied Psychology, 42(1),* 67–71.

Creasey, H., Waite, L.M., Grayson, D.A., Bennett, H.P., Dent, O., & Broe, G.A. (2001). The impact of the neurodegenerative disorders on ageing: An overview of the Sydney Older Persons Study. *Australasian Journal on Ageing, 20 (1),* 11–16.

Crowe, S.F., Dingjan, P., & Helme, R.D. (1997). The neurocognitive basis of word-finding difficulty in Alzheimer's disease. *Australian Psychologist, 32(2),* 114–119.

Farmer, A. (1990). Performance of normal males on the Boston Naming Test and The Word Test. *Aphasiology, 4(3),* 293–296.

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). "Mini-mental state." A practical method for grading the mental state of patients for the clinician. *Journal of Psychiatric Research, 12,* 189–198.

Goodglass, H. (1980). Naming disorders in aphasia and aging. In L.K. Obler & M.L. Albert (Eds.), *Language and Communication in the Elderly.* Lexington: DC Health.

Henderson, L.W., Frank, E.M., Pigatt, T., Abramson, R.K., & Houston, M. (1998). Race, gender, and educational level effects on Boston Naming Test scores. *Aphasiology, 12(10),* 901–911.

Hodges, J.R., Salmon, D.P., & Butters, N. (1991). The nature of the naming deficit in Alzheimer's and Huntington's Disease. *Brain, 114,* 1547–1558.

Howieson, D.B., Holm, L.A., Kaye, J.A., Oken, B.S., & Howieson, J. (1993). Neurologic function in the optimally healthy oldest old: Neuropsychological evaluation. *Neurology, 43,* 1882–1886.

Hughes, C.P., Berg, L., Danziger, W.L., Coben, L.A. & Martin, R.L. (1982). A new clinical scale for the staging of dementia. *British Journal of Psychiatry, 140,* 566–572.

Ivnik, R.J., Malec, J.F., Smith, G.E., Tangalos, E.G. & Petersen, R.C. (1996). Neuropsychological tests' norms above age 55: COWAT, BNT, MAE Token, WRAT-R Reading, AMNART, STROOP, TMT, and JLO. *The Clinical Neuropsychologist, 10(3),* 262–278.

Jacobs, D.M., Sano, M., Dooneief, G., Marder, K., Bell, K.L., & Stern, Y. (1995). Neuropsychological detection and characterisation of preclinical Alzheimer's disease. *Neurology, 45,* 957–962.

Jorm, A.F., Korten, A.E., & Henderson, A.S. (1987). The prevalence of dementia: A quantitative integration of the literature. *Acta Psychiatrica Scandinavica, 76,* 465–479.

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

Knesevich, J., LaBarge, E., & Edwards, D. (1986). Predictive value of the Boston Naming Test in mild senile dementia of the Alzheimer's type. *Psychiatric Research, 19,* 155–161.

Koss, E., Patterson, M.B., Ownby, R., Stuckey, J.C., & Whitehouse, P.J. (1993). Memory evaluation in Alzheimer's disease. Caregivers' appraisals and objective testing. *Archives of Neurology, 50,* 92–97.

LaBarge, E., Balota, D.A., Storandt, M., & Smith, D.S. (1992). An analysis of confrontation naming errors in senile dementia of the Alzheimer type. *Neuropsychology, 6(1),* 77–95.

LaBarge, E., Edwards, D., & Knesevich, J.W. (1986). Performance of normal elderly on the Boston Naming Test. *Brain and Language, 27,* 380–384.

Lichtenberg, P.A., Ross, T., & Christensen, B. (1994). Preliminary normative data on the Boston Naming Test for an older urban population. *The Clinical Neuropsychologist, 8(1),* 109–111.

Mack, W.J., Freed, D.M., Williams, B.W., & Henderson, V.W. (1992). Boston Naming Test: Shortened versions for use in Alzheimer's Disease. *Journal of Gerontology, 47(3),* 154–158.

Mariën, P., Mampaey, E., Vervaet, A., Saerens, J., & De Deyn, P.P. (1998). Normative data for the Boston Naming Test in Native Dutch-speaking Belgian elderly. *Brain and Language, 65,* 447–467.

Morris, J.C., McKeel, D.W. Jr, Storandt, M., Rubin, E.H., Price, J.L., Grant, E.A., Ball, M.J., & Berg, L. (1991). Very mild Alzheimer's disease: Informant-based clinical, psychometric, and pathologic distinction from normal aging. *Neurology,* 469–478.

Neils, J., Baris, J.M., Carter, C., Dell'aira, A.L., Nordloh, S.J., Weiler, E., & Weisiger, B. (1995). Effects of age, education, and living environment on Boston Naming Test performance. *Journal of Speech and Hearing Research, 38,* 1143–1149.

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, 3,* 569–580.

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

Ross, T.P. & Lichtenberg, P.A. (1997). Effects of age and education on neuropsychological test performance: A comparison of normal versus cognitively impaired geriatric medical patients. *Aging, Neuropsychology and Cognition, 4(1),* 74–79.

Ross, T.P., Lichtenberg, P.A., & Christensen, B.K. (1995). Normative data on the Boston Naming Test for elderly adults in a demographically diverse medical sample. *The Clinical Neuropsychologist, 9(4),* 321–325.

Storandt, M., & Hill, R.D. (1989). Very mild senile dementia of the Alzheimer's type. II. Psychometric test performance. *Archives of Neurology, 46,* 383–386.

Tombaugh, T.N., & Hubley, A.M. (1997). The 60-Item Boston Naming Test: Norms for cognitively intact adults aged 25 to 88 years. *Journal of Clinical and Experimental Neuropsychology, 19(6),* 922–932.

Van Gorp, W.G., Satz, P., Kiersch, M.E., & Henry, R. (1986). Normative data on the Boston Naming Test for a group of normal older adults. *Journal of Clinical and Experimental Neuropsychology, 8(6),* 702–705.

Waite, L.M., Broe, G.A., Casey, B., Bennett, H.P., Jorm, A.F., Creasey, H., Cullen, J., & Grayson, D.A. (1998). Screening for dementia using an informant interview. *Aging, Neuropsychology and Cognition, 5(3),* 194–202.

Waite, L.M., Broe, G.A., Creasey, H., Grayson, D.A., Cullen, J., O'Toole, B.I., Edelbrock, D., & Dobson, M. (1997). Neurodegenerative and other chronic degenerative disorders among people in the community aged 75 and over. *Medical Journal of Australia, 167,* 429–432.

Waite, L.M., Broe, G.A., Creasey, H., O'Toole, B., Edelbrock, D., & Grayson, D. (1995). Neuroepidemiology of the "Old Old". *Neurology, 45 (Suppl.4),* A327.

Welch, L.W., 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,* 260–266.

Worrall, L.E., Yiu, E.M-L., Hickson, L.M.H., & Barnett, H.M. (1995). Normative data for the Boston Naming Test for Australian elderly. *Aphasiology, 9(6),* 541–551.