

## Normative Data on a Malay Version of the Boston Naming Test

SANDRA VAN DORT, ETAIN VONG, ROGAYAH A. RAZAK,  
RAHAYU MUSTAFFA KAMAL & HOOI POH MENG

### ABSTRACT

*This paper aims to describe a Malay version of the Boston Naming Test (M-BNT) and its normative data. The M-BNT follows closely the general administration procedures of the original Boston Naming Test (BNT) but is different in terms of item content. A total of 29 items from the original 60 items on the test were deemed culturally and linguistically valid for the Malay population and were thus retained. A total of 41 additional items were added to make a total list of 70 items for pilot testing. These items were first vetted by a panel of experts and then trialed on a sample of 40 Malay adults. Based on the item analysis from the pilot study, the M-BNT was reduced to a 50 item test. This was administered to 230 normal Malay subjects in five age groups (20 – 29 years, 30 – 39 years, 40 – 49 years, 50 – 59 years, and 60 – 69 years), split into two main educational levels (i.e. < 12 years of education, and 12 years or more) and across gender. The Malay subjects were chosen representative of the four major geographical regions in West Malaysia. Initial normative data was computed according to the five age groups and two educational levels. It is hoped that the M-BNT will become a test useful in the identification of patients with an expressive language word-finding disorder.*

*Key words: Speech Pathology, Test Adaptation, Malaysia*

### ABSTRAK

*Kajian ini bertujuan menghuraikan versi Melayu Ujian Penamaan Boston (M-BNT) dan data normatif ujian tersebut. M-BNT menggunakan prosedur pelaksanaan yang serupa dengan Ujian Penamaan Boston yang asal tetapi mempunyai isi kandungan yang berbeza. Sebanyak 29 item daripada ujian asal dikekalkan, kerana didapati sahih dari segi budaya dan bahasa populasi Melayu di Malaysia. Sebanyak 41 item baru disertakan agar sejumlah 70 item digunakan dalam ujian rintis. Pada awalnya, item ujian disemak oleh sekumpulan panel pakar, dan selepas itu, ujian dilakukan ke atas 40 orang Melayu dewasa normal. Berdasarkan analisis item daripada kajian rintis tersebut, M-BNT versi terakhir dibentuk dan terdiri daripada ujian sejumlah*

50 item. Ujian ini kemudian dilakukan ke atas 230 subjek Melayu normal dalam lima kumpulan umur (20 – 29 tahun, 30 – 39 tahun, 40 – 49 tahun, 50 – 59 tahun, dan 60 – 69 tahun), yang dibahagi kepada dua tahap pendidikan (iaitu < 12 tahun pendidikan, dan 12 tahun atau lebih) dan jantina berbeza. Subjek dipilih daripada empat kawasan geografikal utama di Semenanjung Malaysia. Data normatif dihasilkan berdasarkan keputusan M-BNT mengikut lima kumpulan umur dan dua tahap pendidikan. Adalah diharapkan bahawa M-BNT dapat menjadi suatu alat ujian dalam mengenal pasti pesakit yang bermasalah bahasa ekspresif dalam pencarian perkataan .

*Kata kunci: Bidang Patologi Pertuturan, Pengadaptasian Alat Ujian, Malaysia*

## INTRODUCTION

Speech pathologists in Malaysia suffer a great lack of standardized tests that have been normalized on local populations. This hampers our work in documenting and differentially diagnosing communication disorders. A naming test is a component of a speech pathologists' adult language assessment battery in the differential diagnosis of aphasia, dementia and other cognitive-communication disorders. It is critical to use a local test because naming ability is heavily influenced by cultural and linguistic factors. Given all these considerations, we decided to develop a Malay version of the Boston Naming Test. The Boston Naming Test (Kaplan et al. 1983) has been used extensively in the United States and other English speaking countries (Worrall et al. 1995, Heaton et al. 1999) and has been translated and adapted into other languages (Kim & Na 1999, Ardila & Roselli 1989). It consists of a series of 60 black and white line drawings graded according to difficulty level, from pictures having a high degree of imageability and familiarity (eg. "bed," "tree") to pictures of objects which are significantly harder (eg. "protractor," "abacus"). The respondent is asked to name the items and the test stops when the respondent makes seven consecutive naming errors.

In conducting this study, we also looked at the effects of age, education, gender, geographical location and language dominance on naming ability. In Goulet et al.'s (1994) case study review of twenty five studies on the relationship of naming ability and normal aging, there was no single consistent finding regarding the relationship of aging with naming ability. However at least half of the studies listed (e.g. Albert et al. 1988 and Thulliard & Assal 1989) do point to a deterioration in naming ability after 70 years of age in normal adults. The original Boston Naming Test computes mean scores (and standard deviations) for those who have less than 12 years of schooling versus those with greater than 12 years of schooling and finds a significant difference in the performance

with those with better education levels obtaining higher scores. Kim and Na (1999) and Cordell, Senior and Douglas (1998) also come up with a similar pattern. Gender does not contribute significantly to naming abilities (Kim & Na 1998, Radanovic & Mansur 2002).

Although Malaysia is a small country, there are quite a few distinct dialects of Malay spoken among Malay Malaysians (e.g. northern dialect versus East Coast dialect). The main dialect from which standard Malay derives much of its lexicon is Johor Malay; while the dialect that is most dissimilar in terms of lexicon, is the Kelantan dialect. Added to this, many speakers of Malay may also speak Malaysian English particularly in the urban areas. Speakers who are bilingual or trilingual (i.e. use dialect, standard Malay, and/or other local languages) may show different patterns of naming ability compared with monolinguals.

## MATERIALS AND METHODS

### CONSTRUCTION OF THE TEST

Permission to adapt and translate the test was obtained in writing from the test publishers. A total of 29 items from the original test list of 60 items were judged by the researchers (four local speech pathologists and one linguist) as having both cultural as well as linguistic appropriateness and were thus retained. The researchers then made up a list of 41 possible new items for pilot testing, drawing from some unpublished studies on the familiarity and imageability of Malay Nouns and Verbs (Mohd Azmarul Aziz 2004; Rahayu Mustafa Kamal 2001 & Noryantimarlina Abdullah 2001) as well as a picture dictionary of Malay Nouns and Verbs (Ahmad & Yap 1985). The new items attempted to maintain the semantic word category of the original item. For example, we replaced the word “pretzel”, which is an American snack with the word “*ketupat*” (rice cakes) which is a Malaysian food. After the words were selected, line drawings were drawn for all the additional items, ensuring that the size of the line drawings were consistent with the items on the original test. A stimulus cue for each of these items was formulated. A stimulus cue is a phrase that the tester uses to disambiguate pictures that are not recognizable to the respondent. For instance, the stimulus cue for item two on the original test “tree” is “something that grows outdoors”.

The second stage of the adaptation/translation process involved conducting a survey of these 70 pilot items among a panel of experts which comprised an additional seven speech pathologists across Malaysia and one linguist. The panel were asked to name the items, and then to comment on (1) the item appropriateness (2) picture recognizability, and (3) stimulus cue appropriateness. From these comments, five items were redrawn to make them more easily recognizable, two items were withdrawn completely because of the issue of multiple names [*acuan* (biscuit tray); *penyodok* (hoe)] and 25 stimulus

cues were reworded for better clarity. In addition, the panel of experts recommended that dialectal variations should be accepted for ten of these pilot items i.e. (*pensel, sikat, wisel, karipap, penyangkut baju, sempoa, penyiram, beca, burung enggang, pencakar*)

After incorporating the above changes to the pilot items, a total of 68 items were then piloted on a group of 40 individuals selected according to age and gender. The test administration technique was similar to the general procedure on the original BNT. The tester requests the respondent to name the pictures within 15 seconds. If he does not recognize the picture, the stimulus cue is given. If the respondent correctly names the picture, he scores 1 point. If he is incorrect, does not remember the name, or names the item after 15 seconds, he scores 0 points. The total score of the respondent is then summed up. The only difference in this testing procedure from the original BNT was that the test was not stopped even if the respondent made seven consecutive errors as the pictures had not been graded according to difficulty level.

These results were analyzed both qualitatively and quantitatively. In quantitative terms, item discrimination scores (which measures how well the item discriminates between high scorers versus low scorers) and item difficulty scores (which measures the percentage of subjects that were able to name the item) were computed for each test item in order to select the best possible items. In qualitative terms, the researchers noted picture ambiguity/ recognizability difficulties, code-switching and the use of dialectal names. For instance the item 'karipap' (curry puff) was deleted because the redrawn picture still generated many ambiguous responses. Further, the items *kerusi roda* (wheelchair), *raket* (badminton) and *filem* (Kodak) were deleted because a majority of subjects used the Malaysian English/ Colloquial English term (bracketed). Additional dialectal variations were also noted, for example the use of *pedal ayam* for "padlock" among Kelantanese speakers and the use of *sadak* for "sickle" among speakers in Kedah. Based on the results above, a final total of 50 items were deemed to be appropriate to be included in the final version and these were ordered according to item difficulty level from easiest to hardest. It must be noted that this is a change from the original BNT which has 60 items on the test and may have some bearing on the test (refer to Appendix 1 for a list of the words on the BNT-M).

## COLLECTION OF NORMATIVE DATA

230 participants were recruited according to convenience sampling. Subjects in five age groups (20–29 years, 30–39 years, 40–49 years, 50–59 years, and 60–69 years), with an approximately equal number of males ( $n = 119$ ) and females ( $n = 111$ ) were targeted. The subjects came from four major regions in West Malaysia. Participants from Kelantan and Terengganu made up the East Coast sample; participants from Johor and Melaka made up the South sample;

participants from Kedah and Perak made up the North sample, and participants from Kuala Lumpur and Selangor made up the Central sample. All participants were of Malay ethnicity and details about their educational level were recorded. The participants were first interviewed to ensure that they were subjects with no self-reported evidence of neurological problems or a history of speech and language disorders. They were also categorized into dominant speakers of Malay (using Malay in their everyday life > 80% of the time) or speakers whose dominant language is not Malay.

## RESULTS

### GENERAL DESCRIPTIVE ANALYSIS

From a preliminary observation of the data, it was noted that the mean scores of the participants according to gender do not differ significantly. The mean score for males is 46.00 (s.d. = 3.83) and for females is 45.27 (s.d. = 4.06). Table 1, displays the results of the M-BNT according to the stratified groups of education and age level. The highest mean score was obtained by participants 20 -29 years of age, with greater than 12 years of education (M = 49.97 SD = 1.64). On the other hand, participants of 60 - 69 years, with < 12 years of education showed the lowest mean score ( M = 41.76, SD = 5.20 ) .

TABLE 1. BNT Mean test scores according to age group and educational level

Age Group (years)	Education (years)					
	n	< 12 years		12 years or more		
		M	(S.D.)	n	M	(S.D.)
20-29	18	46.89	2.49	31	49.97	1.64
30-39	20	46.55	2.37	26	47.8	2.02
40-49	29	46.06	2.96	16	46.69	2.72
50-59	31	43.90	4.23	18	44.72	3.77
60-69	39	41.76	5.20	2	47.5	2.12

### REGRESSION ANALYSIS

The data was also analyzed by multiple regression analysis, using as regressors the independent variables of age, educational level, geographical location, language dominance, and gender to see if they have any bearing on the results. We categorized location into two categories, that is East Coast location versus the rest of West Malaysia, since a preliminary calculation of means and self-reports by the researchers indicated that only scores of the east coast participants

were different from the other regions. Together these variables accounted for 30.2% of the variation on the BNT scores. The overall relationship was significant ( $F = 20.82, p < 0.0005$ ). Looking at individual variables, age was the most significant ( $\beta = -0.414, p < 0.0005$ ). When age increased, total scores decreased, although it must be noted that in the 3 younger age groups (20-29 years, 30 – 39 years and 40-49 years) the mean scores as listed in Table 1, are very similar and thus the differences between these age groups may not be as significant. Location was also significant ( $\beta = -0.214, p < 0.0005$ ) indicating that east coast subjects scored significantly lower compared with subjects in the other regions. Educational level was mildly significant ( $\beta = 0.183, p < 0.05$ ) indicating that subjects with greater than 12 years education performed slightly better than those with less than 12 years of education. However, language dominance and gender did not have a significant influence on the scores.

GENERATION OF NORMAL SCORES AND CUT-OFF SCORES BASED ON THE  
NORMAL POPULATION RESULTS

Based on the results above, and keeping with the criteria of the original Boston Naming Test, it was decided that normative values and cut-off scores would be generated based on age and educational level. The sample sizes of subjects per cell, based on age and level of education is displayed in Table 2.

TABLE 2. Collection of normative data according to age group and educational level

Age Group	Educational Level		Total
	<12 years	≥ 12 years	
20-29	18	31	49
30-39	20	26	46
40-49	29	16	45
50-59	31	18	49
60-69	39	2	41
Total	123	107	230

There are a similar number of subjects according to age levels, with a range between 41 to 49 participants in each group. There are also a similar number of total participants according to educational level. However, because of the nature of convenience sampling and our nation’s educational history, it is apparent that as age groups get older, there are smaller sample cells of subjects with more than twelve years educational level.

Using the results from Table 2, Table 3 was constructed and indicates cut-off scores for the BNT-M organized according to age and educational level. These cut-off scores are calculated based on general speech pathology testing principles that assume that scores falling two standard deviations from the mean score indicate a communication pathology may be present.

TABLE 3. Cut-off scores for the M-BNT

Age Group (years)	Education (level)	
	< 12 years	≥12 years
20-29	42	45
30-39	42	44
40-49	40	41
50-59	35	37
60-69	31	43

## DISCUSSION

Gender is not a factor that appears to influence naming ability, and this is borne out by data from other countries (Kim & Na 1998; Radanovic & Mansur 2002). In this study, language dominance was also not found to be significant. Only 8.3% of the sample were non dominant speakers of Malay, but this did not appear to have a bearing on their scores. This could point to the fact that Malays who also speak Malaysian English or other languages, are bilingual and able to perform as adequately as monolingual speakers of Malay in the naming test. More studies need to be conducted to investigate this phenomena.

Educational level, however, does have some bearing on naming ability, as predicted by data from other studies (Kim & Na; 1999; Cordell, Senior & Douglas 1998), and this could be related to the low familiarity and imageability of the harder test items.

In our study, age was discovered to be a factor that affected naming ability and this supports the findings of other studies (Albert et al. 1988; Thulliard & Assal 1989) and indicates clearly that the older age groups (0-59 years, 60-69 years) show some deterioration in naming skills.

East coast speakers, show differences in performance from the rest of West Malaysia, and these differences can be related to dialectal variations and cultural differences. For instance, based on personal report from the researchers, items such as “angklung” (a musical instrument) and “tanglung” (a Chinese lantern) are relatively unfamiliar for Malay speakers on the East coast. This means that clinicians who use the M-BNT for East coast speakers should use the cut-off scores and normative values generated in this study with some caution and only as an adjunct to an experienced clinicians’ judgment.

There are some limitations present in the study. Due to the use of the convenience sampling procedure, the number of older participants with greater than 12 years of education is limited in number, and thus the data on these subjects may not be representative. Further, the reduction of length of the M-BNT (i.e. only 50 items on the test), may have some implications on this adapted test, as the original BNT was made up of 60 items.

## CONCLUSION

This was a small study to adapt the original Boston Naming Test into Malay. Initial norms according to age and education, have been collected from a sample of 230 ethnic Malay adult speakers in Peninsula Malaysia and initial cut-off scores have been generated. Much more work is needed to fully develop this test and the norming samples need to be increased and when possible extended to second language speakers of Malay, as well as populations in East Malaysia.

One of the difficulties that any researcher will encounter in creating a naming test in multilingual, multicultural Malaysia, will be the difficulty with dialectal variations of vocabulary, code-switching and code mixing with other languages and multiple names for certain items. Further as is apparent from the data from the east coast, word frequency differences may be different in different regions of Malaysia. These were some of the inherent difficulties of this study, but it is hoped that future researchers will take this into consideration. It is indeed hoped that this initial study will pave the way for more research in this very important area of developing standardized speech-language assessments for our local population.

## ACKNOWLEDGEMENTS

The researchers would like to acknowledge Wan Zaharah bt Abdul Wahid, Nur Syuhada bt Hosni, Yazmin bt Ahmad Rusli, Radziah bt Mohd Fahmy, Norhana bt Abu Seman @ Talib, Wahida bt Mohd Abd Wahab, and Nur Shahrina bt Mohd Zawawi who helped in the collection of data; Prof Dr Jane Maxim at University College London for her initial suggestions and advice; and Universiti Kebangsaan Malaysia for the UKM fundamental grant NN-03102004.

## APPENDIX 1. LIST OF WORDS AND STIMULUS CUES

1. Rumah (sejenis bangunan kediaman)
2. Bunga (tumbuhan di taman)
3. Pokok (sesuatu yang tumbuh di luar rumah)
4. Pensel/ /kālɛ/ (digunakan untuk menulis)
5. Gunting (digunakan untuk memotong)
6. Sikat/ Sisir (digunakan untuk mendandan rambut)
7. Gergaji/ Gaji (digunakan oleh tukang kayu)
8. Unta/ /untɔ/ (sejenis binatang di padang pasir)
9. Keris (sejenis senjata tradisional)
10. Tongkat (alat untuk membantu seseorang berjalan)
11. Katil (sejenis perabot di bilik tidur)
12. Berus gigi (digunakan untuk membersihkan gigi)
13. Buai/ Buaian/ /əndo/ (sejenis peralatan mainan di taman)
14. Bendera (dikibarkan untuk mewakili negara)

15. Sotong (sejenis hidupan laut)
16. Ketupat (sejenis makanan Melayu)
17. Gasing (sejenis permainan tradisional)
18. Beca/ Teksi (sejenis pengangkutan darat)
19. Antena/ Aerial (digunakan untuk mendapatkan siaran TV yang lebih jelas)
20. Khemah (digunakan di luar rumah untuk berlindung)
21. Rakit (sejenis pengangkutan air)
22. Gitar (sejenis alat muzik)
23. Badak sumbu i (sejenis binatang) ii) (badak jenis apa) – jika subjek menjawab ‘badak’ sahaja
24. Belalang (sejenis serangga)
25. Helikopter (digunakan untuk perjalanan udara)
26. Siput (sejenis binatang bercangkerang)
27. Tali gantung i (digunakan untuk hukuman mati) ii) (tali jenis apa) – jika subjek menjawab ‘tali’ sahaja
28. Tupai (sejenis binatang)
29. Sabit/ Sadak/ Penyabit (digunakan untuk memotong rumput)
30. Wisel (digunakan untuk menghasilkan bunyi)
31. Gunung berapi (sejenis gunung)
32. Topeng (digunakan untuk menutup muka)
33. Mangga/ Ibu kunci/ Pedal ayam (digunakan untuk mengunci)
34. Zirafah (sejenis binatang berleher panjang)
35. Burung merak (sejenis burung)
36. Gong (sejenis alat muzik tradisional)
37. Ikan pari/ Tuko (sejenis hidupan laut)
38. Kala jengking/ Kala/ Jengking (sejenis serangga berbisa)
39. Seruling/ Suna (sejenis alat muzik)
40. Pahat/ Pemahat (digunakan oleh tukang kayu)
41. Tanglung (sejenis lampu Cina)
42. Kaktus (sejenis tumbuhan di padang pasir)
43. Kuda laut (sejenis hidupan air)
44. Corong/ Copong/ Calong (digunakan untuk memudahkan penuangan cecair)
45. Piramid (ditemui di Mesir)
46. Kompas (digunakan untuk menentukan arah)
47. Pingat (sejenis anugerah)
48. Burong enggang/ Kenyalang i (sejenis burung) ii) (burung jenis apa) – jika subjek menjawab ‘burung’ sahaja
49. Angklung (sejenis alat muzik tradisional)
50. Sempoa/ Sepua (alat yang digunakan untuk mengira)

## REFERENCES

- Ahmad, S. & Yap, L.K. 1985. *Kamus KBSR Dewan Bergambar*. Kuala Lumpur: Dewan Bahasa & Pustaka.
- Ardila, A. & Rosselli, M. 1989. Neuropsychological characteristics of normal aging. *Dev. Neuropsychol.* 5: 307-320.
- Cordell, S.G. & Douglas, L. 1998. The relationship between word knowledge and verbal fluency in neuropsychological assessment. Australia: Annual Conference of the College of Clinical Neuropsychologists.
- Goulet, P., Ska, B. & Kahn, H.J. 1994. Is there a decline in picture naming with advancing age? *J. Speech Hear. Res.* 37(3): 629-644.
- Heaton, R.K., Avitable, N., Grant, H., Matthews C.G. 1999. Further Cross-Validation of Regression Based Neuropsychological Norms with an update for the Boston Naming Test. *J. Clin. Exp. Neuropsychol.* 99 21 (4): 572-582.
- Kaplan, E., Goodglass, H. & Weintraub, S. 1983. *Boston Naming Test*. USA: Williams and Wilkins.
- Kim, H. & Na, D.L. 1999. Normative data on the Korean version of the Boston Naming Test. *J. Clin. Exp. Neuropsychol.* 21 (1): 127-133.
- Radanovic, M., Mansur, L.L. 2002. Performance of a Brazilian population sample in the Boston Diagnostic Aphasia Examination: a pilot study. *Braz J. Med. Biol. Res.* 35: 305-317.
- Worrall, L.E., Yiu, M-L, Hickson, L.M.H., & Barnett, H.M. 1995. Normative data for the Boston Naming Test for Australian elderly. *Aphasiology* 9: 541-551.

Sandra Van Dort  
Etain Vong  
Rogayah A Razak  
Rahayu Mustaffa Kamal  
Hooi Poh Meng  
Department of Audiology and Speech Sciences  
Faculty of Allied Health Sciences  
Universiti Kebangsaan Malaysia  
Jalan Raja Muda Abdul Aziz  
50300 Kuala Lumpur