NEUROLOGÍA

ORIGINAL ARTICLE

Spanish normative studies in a young adult population (NEURONORMA young adults project): Norms for the Boston Naming Test and the Token Test[☆]

F. Aranciva[a], M. Casals-Coll[a], G. Sánchez-Benavides[a], M. Quintana[a], R.M. Manero[b], T. Rognoni[a], L. Calvo[a], R. Palomo[a], F. Tamayo[a], J. Peña-Casanova[b,*]

[a] *Grupo de Neurología de la Conducta y Demencias, Programa de Neurociencias, Instituto de Investigación Hospital del Mar (IMIM), Barcelona, Spain*
[b] *Sección de Neurología de la Conducta y Demencias, Servicio de Neurología, Hospital del Mar, Barcelona, Spain*

Abstract

Introduction: The Boston Naming Test (BNT) and the Token Test (TT) are frequently used in clinical practice to assess naming and comprehension.
Objective: The aim of this paper is to present normative data for the BNT and for the TT as part of the NEURONORMA young adults project.
Material and methods: A total of 179 Spanish healthy subjects from 18 to 49 years old were evaluated. Tables to convert raw scores to scaled scores are provided. Age- and education-adjusted scores are obtained by applying linear regressions.
Results: The results show an effect of education in both tests, and a minimal effect of age and sex.
Conclusions: The normative data obtained will be useful in the clinical evaluation of young Spanish adults.

Estudios normativos españoles en población adulta joven (Proyecto NEURONORMA jóvenes): normas para el *Boston Naming Test* y el *Token Test*

Resumen

Introducción: El *Boston Naming Test* (BNT) y el *Token Test* (TT) son pruebas ampliamente utilizadas en la práctica clínica para explorar la denominación y la comprensión.
Objetivo: En este artículo se presentan los datos normativos del BNT y el TT obtenidos en el proyecto normativo español NEURONORMA jóvenes.

Material y métodos: La muestra está formada por 179 sujetos sanos de entre 18 y 49 años de edad. Se aportan tablas para convertir las puntuaciones brutas en escalares y tablas con los ajustes pertinentes por edad y escolaridad obtenidas a partir de regresiones lineales.
Resultados: Los resultados obtenidos muestran influencia de la escolaridad en ambos tests, pero un mínimo efecto de la edad y del género.
Conclusiones: Las normas obtenidas aportan datos de gran utilidad clínica para la evaluación de población adulta joven española.

Introduction

The primary objective of the NEURONORMA project (NN) is to collect normative data from the Spanish population using neuropsychological tests frequently employed in assessing cognitive capacities. Normative data have recently been published for subjects older than 50.[1] As an extension of the NN project, the same test battery is currently being normalised in subjects aged 18 to 49 (NEURONORMA young adults project, NNy). The current article presents normative data from two language tests, the Boston Naming Test (BNT) and the Token Test (TT).

Both tests are frequently used in research and clinical practice since they are easy to administer and highly sensitive for detecting language difficulties in a number of different pathologies.[2] Due to their characteristics, both tests may be useful in assessing other cognitive areas such as the integrity of semantic memory (for BNT) or working memory (for TT).[3–6]

The BNT provides information about subjects' ability to name objects depicted in black and white drawings.[7] Kaplan, Goodglass, and Weintraub developed the first experimental version of the BNT with 85 test items. In the standard version, the number of items was reduced to 60.[8] The second edition currently allows subjects to select answers from multiple choice arrays and lets administrators classify the different types of errors.[3] Several short versions have been developed in order to reduce test administration time (for an exhaustive review see Kent and Luszcz[9]).

Performance on the BNT has frequently been linked to sociodemographic factors, such as age, education, intelligence quotient and sex. Several studies state that older subjects present poorer results and that the standard deviation also increases with the age of the group as scores become more dispersed.[7,10–13] However, other authors have not found any significant relationship between age and performance.[14–17]

The influence of education on performance is also controversial. Although there are numerous studies showing influence by this variable,[17–26] other studies do not describe any education effect.[13,15,16,27–29]

Most studies find no significant differences in performance between men and women,[10–13,22,30] but other studies describe a slight advantage for the male population.[11,17,23,24] A study carried out by Randolph et al.[12] suggests that these results are due to a high number of test items on the BNT being male-biased. Other studies detect no influence of sex on performance.[9,13–16,20,22,23,25,28–30]

Numerous studies provide normative data from the BNT in different languages (for a review, see Strauss et al.,[2] Lezak

et al.[31] and Mitrushina et al.[32]). For young adult populations, normative data are available from the 60-item versions in English[7,15–17] and in Spanish.[33,34] A Spanish version of the second edition of the BNT has also been published, and includes normative data for both younger[35] and older Spanish adults.[36,37] Some abbreviated versions in Spanish provide normative data for elderly subjects.[38,39]

The TT is a comprehension test in which the subject must follow verbal directions involving multi-shaped and coloured cards; task complexity increases as the test progresses. Ever since it was first published, the TT[40] has played an important role in neuropsychological examination. Its main purpose is to assess verbal comprehension, including short-term memory for verbal sequences and the ability to understand syntax.[2] There are several versions and modifications of the test, including abbreviated versions (for an exhaustive review see McNeil and Prescott[41]). Non-aphasic persons with at least 4 years of schooling have no difficulty completing the TT, and usually make few or no errors.[31] Studies report a good level of correlation between the TT and other tests of comprehension[42] and verbal production.[43] The TT can be administered to both adults and children.

There are several studies about the influence of sociodemographic factors on TT performance. Some authors have found a slight age effect on scores.[22,37,44] Children's scores have been found to reach adult levels beginning at age 11, which reflects normal development.[45] In adults, Emery[46] compared the performance of a group of subjects aged 75 to 93 and that of a group of subjects aged 30 to 42. Performance was lower in the first group. The ability to complete the task properly is highly conditioned by the subject's years of education.[34,37,44] In elderly adults, it has been observed that years of education have a more pronounced effect than age on subjects' scores.[37] Earlier results indicate that sex has no significant effect on performance.[37,47]

The main objective of this study was to obtain reference data reflecting the performance of Spanish young adults on the BNT and the TT, and adjust scores by sociodemographic variables if necessary.

Material and methods

Subjects

Recruitment methods and sample characteristics have already been described in the project's methodology article. To summarise, we recruited 179 white subjects who had been educated in Spain, regardless of their first language

(for bilingual subjects). The sample was stratified by age and educational level. All subjects were unaffected by cognitive impairment; their scores on the Mini-Mental State Examination[48,49] were ≥24 and their scores on the Memory Impairment Screen[50,51] were ≥4.

Neuropsychological tests

We employed the neuropsychological protocol selected for use within the framework of the NN project.[1] Tests were administered according to the standard procedures described in their manuals.

Boston Naming Test

We used the second edition of the BNT.[3] The test material consisted of 60 drawings representing objects whose names had varying degrees of frequency and familiarity. Item number 19, 'pretzel', was replaced with another food term, '*magdalena*', since people in our area were unlikely to be familiar with the first item. Picture cards were shown one by one to the subjects, who were then asked to name the object shown. We offered clues to subjects who were unable to answer after 20 s, or who provided an incorrect answer. If the error was caused by difficulty recognising the object, subjects were given a semantic clue. If they were able to identify the object, but incapable of recalling the correct word, they were given a phonological clue. If they were still unable to answer after having received the clues, subjects had the option of choosing the target word from a multiple choice list at the end of the test. The maximum score, calculated as the sum of all cards for which a correct answer was given, either spontaneously or after hearing a semantic clue, was 60 points.

Token Test

In this project, we used the abbreviated version of the test,[44] which comprises 36 test items presented in 6 parts. We used 20 plastic cards in 5 different colours (yellow, green, red, black, and white), 2 sizes (large and small) and 2 shapes (square and round). Subjects were asked to carry out a series increasingly complex actions in response to the verbal directions provided by the researcher. The test included 6 parts with a total of 36 commands. On the first 5 parts of the test, if the subject responded incorrectly or did not respond within 5 s, the command was repeated. Subjects who completed the task on the second attempt received half a point instead of 1 point. The highest possible score on the test was 36 points.

Statistical analysis

A standardised statistical analysis was carried out for all the neuropsychological tests included in the project. A brief summary of the procedure is as follows: (a) raw scores were converted to scaled scores (SSs). To do so, we generated an array of cumulative frequencies of raw scores and created 16 percentile ranges. Each of these ranges corresponds to a scaled score between 2 and 18. Raw scores were converted

to SS according to their position within the distribution. This transformation of raw scores to scaled scores produced a normal distribution (mean = 10, standard deviation = 3) to which linear regressions could be applied. (b) SS correlation coefficients (r) and coefficients of determination (R^2) were determined for age, years of education, and sex for each of the tests. (c) The regression coefficient (β) from this analysis was used as the basis for adjusting for sociodemographic factors. The SS was adjusted for age, education, and sex according to the following formula: $SS_{A\&E\&S} = SS - (\beta_1 \times [age - 35] + \beta_2 \times [education - 13] + \beta_3 \times sex)$. The resulting value was truncated to the next lower integer. We adjusted for only those sociodemographic variables accounting for more than 5% of the variance and presenting a significant regression coefficient. For more detailed information about methodology, see the project's methodology article.

Results

Table 1 displays the array of frequencies of raw scores for the entire group aged 18 to 49, with the corresponding scaled scores and percentile ranks. In the table, we use the patient's raw score to find the associated scaled score. The percentile range corresponding to each scaled score provides an assessment of the subject's performance with respect to the reference population. Nevertheless, this scaled score should be adjusted for sociodemographic factors by using the tables provided.

Table 2 displays correlation coefficients (r) and coefficients of determination (R^2) of cognitive variables with age, education, and sex. We found a significant education effect on both tests. This variable accounted for 10.9% of the TT variance and 8.9% of the BNT variance. The effects

Table 1 Scaled and percentile scores of the BNT and the TT.

NSS	Percentile ranges	BNT	TT
2	<1	≤38	≤32
3	1	39	32.5
4	2	40—42	33
5	3—5	43—44	33.5
6	6—10	45—46	—
7	11—18	47—49	34
8	19—28	50	34.5
9	29—40	51—52	35
10	41—59	53—54	35.5
11	60—71	55	—
12	72—81	56	—
13	82—89	—	—
14	90—94	57	—
15	95—97	58	—
16	98	—	—
17	99	—	—
18	>99	59—60	36
Number of subjects		179	179

BNT: Boston Naming Test; SSs: scaled scores; TT: Token Test.

Table 2 Correlation coefficients (r) and coefficients of determination (R^2) of the scaled scores by age, education, and sex.

	Age (years)		Education (years)		Sex	
	r	R^2	r	R^2	r	R^2
BNT	−0.179[a]	0.032	0.278[b]	0.089[c]	−0.162[a]	0.026
TT	−0.088	0.008	0.329[b]	0.109[c]	−0.098	0.010

BNT: Boston Naming Test; TT: Token Test.
[a] Correlation significant at 0.05 level (bilateral).
[b] Correlation significant at 0.01 level (bilateral).
[c] R^2 greater than 0.05.

Table 3 Education-adjusted table for BNT and TT.

	Education (years)												
	8	9	10	11	12	13	14	15	16	17	18	19	20
BNT[a]	+1	+1	0	0	0	0	0	0	0	−1	−1	−1	−1
TT[b]	+2	+1	+1	0	0	0	0	0	−1	−1	−2	−2	−3

BNT: Boston Naming Test; TT: Token Test.
[a] $\beta = 0.259$.
[b] $\beta = 0.471$.

of age and sex on the TT and BNT scores were minimal (<5%), so we did not adjust for those variables.

Adjustments for education were applied by using the regression coefficient obtained from the multiple regression analysis. We have provided a table (Table 3) with the points that should be added to or subtracted from the SS depending on the subject's level of education in order to calculate $SS_{A\&E\&S}$.

Discussion

The main objectives of this study were to collect normative data from 2 language tests administered to Spanish young adults and analyse the influence of these sociodemographic variables on performance. These variables were adjusted when necessary. This study is part of the NEURONORMA young adults project, a normalisation project for cognitive tests in subjects younger than 50.

Boston Naming Test

According to data from this study, the effect of age on the subjects' performance on the BNT is very small. Therefore, scores do not need to be adjusted by age in young adults. These results coincide with those from a previous study of an all-male sample with a similar size and age range.[16] However, they do not coincide with the results found by Kaplan et al.[7] in the sample used to normalise the original test. There are also discrepancies with other studies that include subjects with a wider age range. By including elderly subjects, these studies have shown a clear age effect.[10–13] Some authors state that the influence of age may be significant beginning at 60 years[52] or as late as 80 years.[23] Our data reveal a moderate education effect on performance on the

BNT. These findings coincide with those from several earlier studies,[17–26] including studies conducted with Spanish speakers.[33] The effect of sex on the performance of the tasks was minimal. This fact coincides with other earlier studies that did not find a relationship between naming ability and sex.[10–13,22,30]

In general, the results observed in our study resembled those from other studies with similar samples of young adults in which an education effect was found, but age and sex effects were not.[34,35]

Compared to the NN project study carried out in subjects older than 50,[37] the results of the current study show differences in the effect of age on performance, probably due to the effect of ageing on visual confrontation naming. The data from both normative studies support the hypothesis that naming ability starts to decline in subjects older than 50. Results for age and sex coincide with the findings in older adults. The relationship between these variables and the BNT therefore remains stable over time.

Token Test

The results from the current study showed a significant education effect on TT scores, as do other prior studies.[33,43] We did not find an age effect on performance, probably because the task's ceiling effect occurs during adolescence.[45] We also found no relationship between performance and sex, which concurs with previous studies.[47]

As we observed for the BNT scores, educational level had a greater effect than age on TT scores. The pattern of the age and sex effects is similar in both tests, which was also true for the normative study in subjects older than 50[37]; in that study, we found that age but not sex affected performance.

Conclusions

The current study provides normative data that will be useful for evaluating language disorders in young adults. We include scaled score conversion tables and education adjustment tables, as the latter is the only sociodemographic variable with a significant influence on performance. We should highlight that prior to this study, no normative data were available for the TT from this age sector of the Spanish population. Our study therefore provides the objective reference data needed to evaluate language in pathologies that are prevalent in young adults, such as cranial trauma or epilepsy. Since these data form part of a larger normalisation study, they permit us to compare performances between tests measuring different cognitive areas. As a result, we can create a complete profile that aids in describing and diagnosing subjects.

The limitations of the complete normalisation project are described in the project's methodology article. Regarding the specific limitations of this study, we should mention the distribution of the variables, especially in the TT; their behaviour is polarised due to the relative simplicity of the items. This creates a ceiling effect for the scores from healthy subjects. In spite of these limitations, we used homogeneous statistical treatment during the entire project to facilitate comparisons between performances on different tests in the NN battery.

The normative data from the language tests presented in this article may be extremely useful for examining young Spanish patients with language disorders.

Conflict of interest

The authors have no conflicts of interest to declare.

References

1. Peña-Casanova J, Blesa R, Aguilar M, Gramunt-Fombuena N, Gómez-Ansón B, Oliva R, et al. Spanish Multicenter Normative Studies (NEURONORMA project): methods and sample characteristics. Arch Clin Neuropsychol. 2009;24: 307—19.
2. Strauss E, Sherman E, Spreen O. A compendium of neuropsychological tests: administration, norms, and commentary. Oxford: Oxford University Press; 2006.
3. Kaplan E, Goodglass H, Weintraub S. Boston Naming Test. 2nd ed. Philadelphia: Lippincott Williams & Wilkins; 2001.
4. Storandt M, Hill R. Very mild senile dementia of the Alzheimer type II: psychometric test performance. Arch Neurol. 1989;46:383—6.
5. Beatty WW, Monson N. Lexical processing in Parkinson's disease and multiple sclerosis. J Geriatr Psychiatry Neurol. 1989;2:145—52.
6. Snitz BE, Unverzagt FW, Chang CC, Bilt JV, Gao S, Saxton J, et al. Effects of age, gender, education and race on two tests of language ability in community-based older adults. Int Psychogeriatr. 2009;21:1051—62.
7. Kaplan E, Goodglass H, Weintraub S. Boston Naming Test. Philadelphia: Lea & Febiger; 1983.
8. Kaplan E, Goodglass H, Weintraub S. Boston Naming Test: experimental edition. Boston: Kaplan & Goodglass; 1978.
9. Kent PS, Luszcz MA. A review of the Boston Naming Test and multiple-occasion normative data for older adults on 15-item versions. Clin Neuropsychol. 2002;16:555—74.
10. Zec RF, Burkett NR, Markwell SJ, Larsen DL. A cross-sectional study of the effects of age, education, and gender on the Boston Naming Test. Clin Neuropsychol. 2007;21:587—616.
11. Marien P, Mampaey E, Vervaet A, Saerens J, De Deyn PP. Normative data for the Boston naming test in native Dutch-speaking Belgian elderly. Brain Lang. 1998;65:447—67.
12. Randolph C, Lansing AE, Ivnik RJ, Cullum CM, Hermann BP. Determinants of confrontation naming performance. Arch Clin Neuropsychol. 1999;14:489—96.
13. Fastenau PS, Denburg NL, Mauer BA. Parallel short forms for the Boston Naming Test: psychometric properties and norms for older adults. J Clin Exp Neuropsychol. 1998;20:828—34.
14. Van Gorp WG, Satz P, Kiersch ME, Henry R. Normative data on the Boston naming test for a group of normal older adults. J Clin Exp Neuropsychol. 1986;8:702—5.
15. Heaton RK, Avitable N, Grant I, Matthews CG. Further crossvalidation of regression-based neuropsychological norms with an update for the Boston naming test. J Clin Exp Neuropsychol. 1999;21:571—82.
16. Farmer A. Performance of normal males on the Boston naming test and the word test. Aphasiology. 1990;4:293—6.
17. Tombaugh TN, Hubley AM. The 60-item Boston Naming Test: norms for cognitively intact adults aged 25 to 88 years. J Clin Exp Neuropsychol. 1997;19:922—32.
18. Nicholas LE, Brookshire RH, MacLennan DL, Schumacher JG, Porrazzo SA. Revised administration and scoring procedures for the Boston Naming Test and norms for non-brain-damaged adults. Aphasiology. 1989;3:569—80.
19. Neils J, Baris JM, Carter C, Dell'aira AL, Nordloh SJ, Weiler E, et al. Effects of age, education, and living environment on Boston naming test performance. J Speech Lang Hear Res. 1995;38:1143—9.
20. Ross TP, Lichtenberg PA, Christensen BK. Normative data on the Boston naming test for elderly adults in a demographically diverse medical sample. Clin Neuropsychol. 1995;9:321—5.
21. Worrall LE, Yiu EML, Hickson LMH, Barnett HM. Normative data for the Boston naming test for Australian elderly. Aphasiology. 1995;9:541—51.
22. Ivnik RJ, Malec J, Smith G, Tangalos E, Petersen R. Neuropsychological testing norms above age 55: COWAT, BNT, MAE, TOKEN, WRAT-R Reading, AMNART, Stroop, TMT, and JLO. Clin Neuropsychol. 1996;10:262—78.
23. Welch LW, Doineau D, Johnson S, King D. Educational and gender normative data for the Boston Naming Test in a group of older adults. Brain Lang. 1996;53:260—6.
24. Lansing AE, Ivnik RJ, Cullum CM, Randolph C. An empirically derived short form of the Boston naming test. Arch Clin Neuropsychol. 1999;14:481—7.
25. Saxton J, Ratcliff G, Munro CA, Coffey CE, Becker JE, Fried L, et al. Normative data on the Boston naming test and two equivalent 30-item short-forms. Clin Neuropsychol. 2000;14:526—34.
26. Piguet O, Millar JL, Bennett HP, Lye TC, Creasey H, Broe GA. Boston Naming Test: normative data for older Australians. Brain Impair. 2001;2:131—9.
27. Mitrushina M, Satz P. Differential decline of specific memory components in normal aging. Brain Dysfunct. 1989;2:330—5.
28. Howieson DB, Holm MA, Kaye JA, Oken BS, Howieson J. Neurologic function in the optimally healthy oldest old: neuropsychological evaluation. Neurology. 1993;43:1882—6.
29. Lichtenberg PA, Ross T, Christensen B. Preliminary normative data on the Boston Naming Test for an older urban population. Clin Neuropsychol. 1994;8:109—11.
30. Cruice MN, Worral LE, Hickson LM. Boston Naming Test results for healthy older Australians: a longitudinal and cross-sectional study. Aphasiology. 2000;14:143—55.

31. Lezak MD, Howieson DB, Loring DW. Neuropsychological assessment. 4th ed. New York: Oxford University Press; 2004.

32. Mitrushina M, Boone KB, Razani J, D'Elia LF. Handbook of normative data for neuropsychological assessment. 2nd ed. New York: Oxford University Press; 2005.

33. Allegri R, Villavicencio A, Taragano F, Rymberg S, Mangone C, Baumann D. Spanish Boston naming test norms. Clin Neuropsychol. 1997;11:416—20.

34. Ardila R, Roselli M, Puente A. Neuropsychological evaluation of the Spanish speaker. New York: Plenum; 1994.

35. Quiñones-Úbeda S, Peña-Casanova J, Böhm P, Gramunt-Fombuena N, Comas L. Preliminary normative data for the second edition of the Boston Naming Test for young Spanish adults. Neurología. 2004;19:248—53.

36. Rami L, Serradell M, Bosch B, Caprile C, Sekler A, Villar A, et al. Normative data for the Boston Naming Test and the Pyramids and Palm Trees Test in the elderly Spanish population. J Clin Exp Neuropsychol. 2008;30:1—6.

37. Peña-Casanova J, Quiñones-Ubeda S, Gramunt-Fombuena N, Aguilar M, Casas L, Molinuevo JL, et al. Spanish Multicenter Normative Studies (NEURONORMA Project): norms for Boston naming test and token test. Arch Clin Neuropsychol. 2009;24:343—54.

38. Calero MD, Arnedo ML, Navarro E, Ruiz-Pedrosa M, Carnero C. 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. 2002;57:187—91.

39. Serrano C, Allegri RF, Drake M, Butman J, Harris P, Nagle C, et al. Versión abreviada en español del test de denominación de Boston: su utilidad en el diagnóstico diferencial de la enfermedad de Alzheimer. Rev Neurol. 2001;33:624—7.

40. De Renzi E, Vignolo L. The token test: a sensitive test to detect receptive disturbances in aphasics. Brain. 1962;85:665—78.

41. McNeil M, Prescott T. Revised token test. Austin: Pro-Ed Inc.; 1978.

42. Morley GK, Lundgren S, Haxby J. Comparison and clinical applicability of auditory comprehension scores on the behavioral neurology deficit evaluation, Boston diagnostic aphasia examination, porch index of communicative ability and token test. J Clin Neuropsychol. 1979;1:249—58.

43. Gutbrod K, Mager B, Meier E, Cohen R. Cognitive processing of tokens and their description in aphasia. Brain Lang. 1985;25:37—51.

44. De Renzi E, Faglioni P. Normative data and screening power of a shortened version of the Token Test. Cortex. 1978;14: 41—9.

45. Rich JB. Pictorial and verbal implicit recognition memory in aging and Alzheimer's disease: a transfer-appropriate processing account. Ph.D. Dissertation, University of Victoria, Victoria, B.C, Canada; 1993.

46. Emery O. Linguistic decrement in normal aging. Lang Commun. 1986;6:47—64.

47. Sarno MT, Buonaguro A, Levita E. Gender and recovery from aphasia after stroke. J Nerv Ment Dis. 1985;173:605—9.

48. Blesa R, Pujol M, Aguilar M, Santacruz P, Bertrán-Serra I, Hernández G, et al. Clinical validity of the mini-mental state for Spanish speaking communities. Neuropsychologia. 2001;39:1150—7.

49. Folstein MF, Folstein SE, McHugh PR. Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res. 1975;12:189—98.

50. Böhm P, Peña-Casanova J, Manero RM, Terrón C, Gramunt N, Badenas S. Preliminary data on discriminative validity and normative data for a Spanish version of the Memory Impairment Screen (MIS). Int Psychogeriatr. 2003;15:249.

51. Buschke H, Kuslansky G, Katz M, Stewart WF, Sliwinski MJ, Eckholdt HM, et al. Screening for dementia with the Memory Impairment Screen. Neurology. 1999;52:231—8.

52. Goodglass H. Disorders of naming following brain injury. Am Sci. 1980;68:647—55.