

Adaptation and Norm Determination Study of the Boston Naming Test for Healthy Turkish Elderly

Ayten EKİNCİ SOYLU¹ , Banu CANGÖZ² 

¹Department of Neurology, Erciyes University Faculty of Medicine, Kayseri, Turkey

²Department of Psychology, Hacettepe University Faculty of Letters, Ankara, Turkey

ABSTRACT

Introduction: The main purpose of this research is develop the Turkish version of the BNT long form (consist of 60 items) [BNT-60 (TR)] and to determine the normative data for Turkish healthy geriatric population. BNT is a neuropsychological test which was widely used to measure naming disorders associated with a variety of neuropathological events. This research consists of two stages. In the stage of pilot study, adaptation of test was completed and BNT-60 (TR) version was developed; and in the stage of normative study, normative data was collected and norm determination was completed.

Methods: Ninety healthy and volunteer elderly were participated in pilot study and 317 were in normative study. Three screening tests called Montreal Cognitive Assessment (MOCA), Functional Activities Questionnaire (FAQ) and Geriatric Depression Scale (GDS) were administered for participant selection. BNT-60 (TR) was applied to participants who meet the inclusion criteria.

Results: According to 5 (age) x 2 (gender) x 3 (education) factorial ANOVA results, main effects of age and education level on BNT-60 (TR) total score were found statistically significant. Then according to MANOVA results, main effects of age and education level on BNT-60 (TR) sub-scores were found statistically significant. On the other hand, main effect of gender was not significant on BNT-60 (TR) scores. The age and BNT-60 (TR) total score were negatively correlated. This results consistent with other normative studies of BNT in the literature.

Conclusion: Finally, BNT-60 (TR) is adopted for Turkish culture, determined normative data and a test which is evaluating naming ability of the older adults was put into use.

Keywords: Boston Naming Test Turkish Version, adaptation, norm determination, neuropsychological tests, object naming

Cite this article as: Ekinci Soylu A, Cangöz B. Adaptation and Norm Determination Study of the Boston Naming Test for Healthy Turkish Elderly. Arch Neuropsychiatry 2018;55:341-348. https://doi.org/10.5152/npa.2017.19331

INTRODUCTION

The Boston Naming Test (BNT) is the best known neuropsychological test used widely for evaluating linguistic ability which includes object naming and word retrieval (1). The BNT has been used in many clinics and scientific studies for carrying out neuropsychological evaluations on children, adults and elderly individuals with different clinical pathologies such as communication disorder, aphasia, dementia or brain lesion. It is accepted that naming performance measured via BNT has an important function for distinguishing cognitive decline due to healthy aging and pathological function disorders in dementia (2,3). The first experimental version of the BNT is comprised of 85 items and the long version with 60 item that is currently used has been prepared by selecting the items from among the 85 that were stronger with regard to psychometric properties (4,5).

The BNT has been adopted to the languages of different countries such as Portugal (1), Spain (6,7), Belgium (8), Korea (9), Australia (10,11), Sweden (12), Brazil (13), New Zealand (14), Greece (15) and France (16) and norms specific to related cultures have been determined.

It is known that the variables of age, gender and education have statistically significant effects on BNT performance independent of

neurological states (1). Even though there are conflicting findings related with the change with respect to age in BNT scores, the dominant opinion in majority of the studies is that naming ability decreases with increasing age (5,16,17). Findings on whether BNT performance differs with respect to gender or not are not sufficient (15). Whereas a general consensus has been attained regarding the fact that education level is an effective variable for naming ability (5,13,16,18). Whereas test results are interpreted especially for diagnosis purposes, the probable effects of age and education level should be kept in mind (19). Otherwise, test results can be misinterpreted (20). Scores acquired from BNT will be valid only in the existence of normative data for evaluating both clinical and healthy sample groups (21). It should be taken into consideration that cultural factors may be effective in interpreting BNT scores in addition to the aforementioned demographic variables. It is not sufficient to translate the BNT only to the language it will be used in when adapting BNT (2,12). In this regard; it is important to examine within the scope of adaptation and norm determination studies whether the test items are suited to the culture in question or not.

According to the population projections put forth by the Turkish Statistics Institute (TUIK) (22), it is estimated that the elderly (65+ age) population

ratio of Turkey which was 8% in 2014 will reach 10,2% in 2023 and that it will be included among countries with a "very old" population according to the classification by United Nations. In this regard, dementia constitutes a significant public health issue with regard to the planning of future healthcare services in societies in which elderly population is increasing (23). According to data put forth by TUIK (22), whereas the population ratio that passed away as a result of Alzheimer type dementia (ATD) which was the most frequent reason for dementia in Turkey in 2011 was 2,9%, it increased to 3,4% in 2012 and to 3,6% in 2013. However, there is no neuropsychological test for objectively measuring the linguistic ability of the elderly population in our country despite this sociodemographic transformation. Recent studies indicate that the BNT is a reliable tool for distinguishing the effect of healthy aging or neurodegenerative diseases such as ATD on naming ability (24,25). Hence, it is important to determine the norm values of the BNT for the elderly group in our country so that it can be used both in scientific studies (to be able to measure the linguistic abilities of healthy elderly individuals in a reliable manner which is one of the most important cognitive functions) and in applied scientific studies (for the diagnosis/definitive diagnosis/early diagnosis of dementias, for the diagnosis of brain damage, for the diagnosis/definitive diagnosis of aphasia). In this regard, BNT is a test with international recognition, literature richness and thus that has comparability. In addition, cognitive deteriorations that can be measured via the test cannot be hidden by compensating mechanisms as easily as in other tests (21). There is an R&D study that was previously carried out in our country for the BNT on a Turkish sample group between the ages of 20-79 (26). However, the adaptation study carried out within the concept of this study had various limitations regarding especially the evaluation of elderly individuals and dementia patients since it was carried out only on 57 participants, that the elderly group is not covered sufficiently (no data has been acquired for the group aged 79 and above) and that various changes have been made in participant selection criteria along with the number of original items that are not in accordance with the original test (details on the reasons have been presented in the "Discussion" section of the article).

In short; the main objective of this study was to carry out an adaptation and norm determination study for the BNT-60 on a sample group comprised of healthy Turkish elderly people (60+ age) whose native language is Turkish; generate the 60 item Turkish version [BNT-60 (TR)] of the BNT and to determine the norm values according to age, gender, education level. Thus, the objective was to present the BNT-60 (TR) to the use of related individuals as a tool that measures object naming which is one of the fundamental linguistic abilities both for clinical applications and pure scientific studies.

METHODS

Pilot Study

First of all, Turkish-English, English-Turkish two-way translations of the application and scoring instructions were carried out. A total of 120 volunteers whose native language is Turkish participated in the pilot study with 30 healthy adults (age range of 19-59) and 90 healthy elderly adults (age range of 60-81). The adult group consisted of 18 females (60%), 12 males (40%); whereas the elderly adult group consisted of 42 females (46.7%) and 48 male (53.3%) participants. When the education levels are considered, there is 1 primary school graduate in the adult group, 7 secondary-high school graduates and 22 university/post-graduate alumni; whereas there are 29 primary school graduates, 29 secondary-high school graduates and 32 university/post-graduate alumni in the elderly adult group.

A Picture Evaluation Form was prepared in order to determine the original BNT-60 items that do not conform to the inclusion criteria [(1. Words translated into Turkish from English with two consecutive consonants

(e.g. *tripod*; In Turkish "*tripot*"), 2. Words the Turkish equivalent of which corresponds to two words (e.g. *toothbrush*; In Turkish "*diş fırçası*"), 3. Words with synonyms in Turkish (e.g. *acorn*; In Turkish "*palamut*" a fish species / a tree species), 4. Objects that are foreign to our culture (e.g. *igloo*; In Turkish "*iglo*"), 5. Monosyllabic words (e.g. *dart*; In Turkish "*dart*"), 6. Compound words (e.g. *seahorse*; In Turkish "*denizati*"), 7. Objects that can have more than one name (e.g. *bed*, In Turkish "*yatak*" "*döşek*" / "*karyola*"), 8. Words with a plural suffix (included in multiple-choice items) (e.g. *beads*; In Turkish "*boncuklar*"), 9. Verbs (included in multiple-choice items) (e.g. *write*; In Turkish "*yazmak*") within the scope of the study carried out for the adaptation to our culture/language. In addition to the original BNT-60 items, the Picture Evaluation Form consists of a total of 105 black and white hand-drawn pictures determined by the researchers including 45 alternative items. Those from among the original BNT-60 items which do not fit the aforementioned criteria were excluded as a result of the feedbacks received and the results of the pilot study data and were replaced by alternative items. Since original test items are listed arranged according to difficulty, the new items in the TR version were rearranged depending on difficulty in accordance with the *Item Difficulty Analysis* results specified in the "Results" section. In conclusion, the BNT-60 (TR) is comprised of 31 original and 29 new items for a total of 60 items. Final BNT-60 (TR) version comprised of 60 items was generated as a result of the pilot study with 20 items each in three different difficulty categories of easy, moderate and hard.

Norm Determination Study Sample Group

The participants of the study consisted of healthy elderly individuals working at and/or retired from the public and/or private institutions and establishments in the cities of Ankara and Kayseri (in Turkey) who have not participated in the pilot study and whose native language is Turkish. The study was carried out during November 2015-May 2016 with a total of 317 volunteer healthy elderly participants in the 60-92 age range ($\bar{X}=71.08$, $SD=7.83$); with 149 females (47%) and 168 males (53%) from the 60-64, 65-69, 70-74, 75-79, 80 and above age groups. 107 of the participants from the sample group were primary school graduates, whereas 107 were secondary-high school graduates and 103 were university/post graduate alumni.

Three screening tests and/or scales were applied to determine whether the participants are elderly adults with healthy cognitive functions. These were: Montreal Cognitive Assessment (MOCA) (27,28), Functional Activities Questionnaire (FAQ) (29,30) and Geriatrics Depression Scale (GDS) (31,32). Studies were carried out for the adaptation to our culture of the screening tests and/or scales used in the study, for norm determination and/or validity and reliability. Participants who received a score of below 21 from MOCA, those in the 50-69 age group who received scores of 5 or above from two or more activities in FAQ, those in the 70 years old and above age group who received scores of 9 or above from three or more activities or who received scores of 14 and above from GDS along with participants who have a history of neurological or psychiatric disorder and who use drugs for these diseases were not included in the study. Demographic characteristics along with the mean and standard deviation values for the screening test scores have been given in Table 1.

Data Acquisition Tools

The following three screening tests were used for participant selection prior to the norm determination study for the BNT-60 (TR).

Montreal Cognitive Assessment (MOCA): The scale developed by Nasreddine et.al. (27) evaluates different cognitive functions comprised of attention and concentration, executive functions, memory, language, visual-spatial skills, abstract thinking and orientation. MOCA is a screening

scale that has been developed especially for evaluating the early stages of cognitive disorder. The Turkish adaptation and reliability study for MOCA was carried out by Selekler, Cangöz and Uluç (28). The lowest score that can be obtained from the scale is 0, the maximum score is 30. Break point for MOCA was set as 21. Scores of 21 and above are evaluated in the normal boundary.

Functional Activities Questionnaire (FAQ): The questionnaire developed by Pfeffer et.al. (29) evaluates the performance of individuals aged 50 and above via in daily activities via a 10 items. The survey is applied on the person himself/herself and/or first degree relatives. The adaptation of the questionnaire for Turkish sample group and norm determination study was carried out by Selekler, Cangöz and Karakoç (30). Each item is scored between 0-3. The lowest score that can be obtained from the questionnaire is 0, whereas the maximum score is 30. Receiving a score of '5 or above' from two or more activities in the 50-69 age group, receiving a score of '9 or above' from three or more activities in the age group of 70 and above indicates that there is a disorder in cognitive activities and dependency to others.

Geriatric Depression Scale (GDS): The scale developed by Yesavage et.al. (31) evaluates the level and intensity of only depressive symptoms for individuals aged 60 and above without questions on sleep disorders, sexual function disorder and problems related with somatic complaints. GDS is comprised of a total of 30 questions the responses of which are "yes" or "no". Every response in favor of depression is evaluated as 1 point, whereas other responses are evaluated as 0. The Turkish adaptation, validity and reliability study was carried out by Ertan and Eker (32). The lowest score that can be obtained from the scale is 0, whereas the maximum score is 30. A score of 14 and above indicates the existence of depression.

Boston Naming Test (BNT-60): The original test developed by Kaplan, Goodglass and Weintraub (33) is a precision measurement tool used frequently for evaluating cognitive disorder, dementia and aphasia. The BNT-60 is the longest version of the test with 60 items. Even though it varies according to the long or short version of the applied test, the application generally takes between 10-20 minutes and the BNT-60 is presented to the participant as a booklet of 60 different stimulant cards with black-white drawings (e.g. *scissors, racket, abacus*; sequentially In Turkish "*makas*", "*raket*", "*abaküs*") the difficulties of which continue to increase. Original test items are comprised of hand-drawn objects which are not prone to uncertainty and which do not have other alternative names. A time of 20 seconds is given to the participant for naming each item. If the participant cannot name (does not respond to) an item (e.g. *harmonica*; In Turkish "*mızıka*") or names it as something else (gives a wrong response) a semantic (lexical) cue (e.g. *musical instrument*; In

Turkish "*müzik aleti*") is given. If the participant cannot name the object in the picture despite the semantic cue provided, this time the first syllable (e.g. *har-*; In Turkish *mi-*) is provided as a phonemic (vocalic) cue. Finally, if the correct response could not be given after the phonemic cue, the first item that could not be named is brought up again and the multiple choice form (they are items with 4 choices located at the back of each stimulus card) (e.g. *flute, harmonica, chocolate, melodica*; sequentially in Turkish "*flüt*", "*mızıka*", "*çikolata*", "*melodika*") is presented. The participant/patient is asked to show or say out loud which one of the 4 items he/she thinks the right item is. The scores are recorded to the scoring form of the test by the researcher. There are two sections in the scoring form which are "summary of the scores" and "scoring of the error types". The total score is comprised of spontaneously responded correct responses and/or correct responses given after cues (only semantic cues). Error types are used to classify wrong responses. It is scored according to five error categories which are phonological (more than half of the phonetic of the target word is preserved), verbal (semantically related with the target word), neological (an overlap of less than 50% with the target word phonetic), multi-word (using more than one word for naming) and perceptual (erroneous perception of the picture). The lowest score that can be obtained from the BNT-60 is 0, whereas the highest score is 60.

Boston Naming Test 60-Item Turkish Version [BNT-60 (TR)]: A total of 29 original test items were replaced by new items as a result of a comprehensive pilot study; accordingly, the original item difficulty ordering was also changed. Final version of the BNT-60 (TR) as a result of the pilot study was printed in exactly the same format as the original test material (as square spiraled booklet with dimensions of 16 x 16 cm), it was translated into our language in accordance with the changes in the standard application and scoring directives and was made ready for use in norm determination study.

Procedure

Hacettepe University Senate Ethical Commission Approval was taken for the study dated 06.04.2015 and numbered 76000869/433-1060. The original test material set of the BNT-60 (test booklet and application/scoring directive) were purchased by the researchers from PRO-ED Inc., USA (www.proedinc.com) company and there is no legal barrier for the its use in research studies and the publication of the research study. Norm determination stage for elderly sample group was started following the completion of the BNT-60 (TR) adaptation study.

The participants were informed about the objective prior to starting the study and informed consent forms were taken. The participants then filled the demographic information form. After this stage, three screening tests (MOCA, FAQ and GDS) were applied in random order and in a single session to each participant who fulfilled the inclusion criteria. Whereas

Table 1. Summary Table in Relating to Demographic Characteristics along with Mean and Standart Deviations of Screening Test Scores of the Participants (N= 317)

Demographic Characteristics	
Age	\bar{X} =71.08 (7.83) Age range=60-92
Gender	Female=149 (47%), Male=168 (53%)
Education level	Primary (0-5)=107 (33.8%), Secondary-high (6-11)=107 (33.8%), University-post (12 and over)=103 (32.4%)
Native language	Turkish=317 (100%)
Hand preference	Right hand=290 (91.5%), Left hand=7 (2.2%), Ambidextrous=20 (6.3%)
Sinistrality in first degree relatives	Exist=54 (%17), Absent=263 (%83)
Mean and Standard Deviations of Screening Test Scores	
MOCA score	23.57 (2.25)
FAQ score	0.15 (0.46)
GDS score	3.54 (3.41)

MOCA: Montreal Montreal Cognitive Assessment, FAQ: Functional Activities Questionnaire, GDS: Geriatric Depression Scale

the BNT-60 (TR) was applied during another session in order to prevent the effects of fatigue. Data acquisition process was applied individually in environments where the participants who make up the sample group are located (public and private institutions and/or establishments, hospitals, clubhouses, associations, houses, etc.).

Pictures in the test booklet comprised of picture cards with dimensions of 16 x 16 cm were arranged in order of increasing difficulty during norm determination study. In accordance with the BNT-60 (TR) standard application instruction, the BNT-60 (TR) is started by presenting Item 30 (crocodile; In Turkish "timsah") to each participant and if this item is answered correctly, all unapplied items were accepted as correctly answered and a score of 29 is given automatically (the participant is accepted as having answered all previous items correctly). The test is continued backwards as long as he/she does not give a wrong answer before Item 38 (rhinoceros; In Turkish: "gergedan"). In case of such an erroneous response, the test goes back to Item 29 (swing; In Turkish "salinca") and the test items are worked backwards until eight pictures in a row are answered correctly. The items before this new starting point are automatically included in the score. The test is stopped and is not continued after eight consecutive mistakes. The reactions of the participant are recorded by the applier during the application. This study was in compliance with the original standard application and scoring directives for the BNT-60.

The approximate application time of the BNT-60 (TR) when applied individually is about 15 minutes. The total application time of the study is about 45 minutes excluding the resting time between the two sessions (screening scales and BNT-60 (TR) applications).

Statistical Analysis

Responses to the Picture Evaluation Form were subject to frequency analysis in order to determine the TR version items within the scope of the pilot study after which *Item Difficulty Analysis* was carried out for examining the difficulty (p) and distinctiveness (r) properties of the test items.

Among the variables examined within the scope of the norm determination study, the age variable has 5 (60-64, 65-69, 70-74, 75-79, 80 and above), the gender variable has 2 (female and male) and the education level variable has 3 (0-5 years, 6-11 years, 12 years and above) levels. All independent variables were manipulated as between groups. Whereas the dependent measure is the BNT-60 (TR) scores. For BNT-60 (TR) total score, 5x2x3 analysis of variance (ANOVA) and for the 7 sub-scores and 3 error scores obtained from BNT-60 (TR) 5x2x3 multivariate analysis of variance (MANOVA) was applied. *Post hoc* analyses (*Bonferroni Correction*) and paired comparisons were carried out for determining the source of the main and interaction effects determined to be statistically significant as a result of ANOVA and MANOVA; that is, for determining the level combinations that cause it. Finally, *Pearson Product Moments Correlation Analysis* was carried out for putting forth the relations between all dependent variables (total score, number of spontaneously given correct responses, number of semantic cues given, number of correct responses to semantic cues, number of phonemic cues given, number of correct responses to phonemic cues, number of multiple choice and number of correct responses to multiple choice) and age variable as well as among themselves. Data analysis was carried out via *Statistical Package Software for the Social Sciences 20.0* (IBM SPSS Statistics for Windows, Version 20.0; IBM Corp; Armonk, NY, ABD).

RESULTS

Pilot Study Results: Item difficulty (p) and item distinctiveness (r) indexes applied for objective tests within the scope of the *Item Difficulty Analysis*

applied for the BNT-60 (TR) version items based on the responses of the 120 healthy participants to the Picture Evaluation Form were calculated for each item separately. Difficulty index puts forth how correct the responses to the items are, whereas distinctiveness index shows how successful the items are in distinguishing the desired property (34). The functionality of difficulty, distinctiveness and distractors were considered together in order to make a healthy decision regarding item selections and the items were interpreted according to the values they got. In this regard, first the difficulty index (p) values were taken into consideration when arranging the items in order and the percentages of the correct responses were taken into consideration if the items have the same index. Item difficulty index varies between 0 and 1, whereas item distinctiveness index varies between -1 and +1. It was determined as a result of the pilot study that the difficulty index of the original items varied between 0 and 1, distinctiveness index varied between -0.25 and 0.92; the difficulty index of the alternative items varied between 0.06 and 1 and that the distinctiveness index varied between -0.33 and 0.88.

Norm Determination Study Results: Analyses related with the norm determination study were carried out based on the BNT-60 (TR) scores of 317 healthy elderly participants. Prior to the analyses, the data were subject to operations such as determination of whether there are any lost values or not, determination of the outliers and/or assigning new values in their stead. Responses to the continuous variables in the raw data were transformed into z values, the value of ± 3.29 was taken as basis when determining the outliers and the values outside this interval were accepted as outliers (35). In this regard, required corrections were made for the scores of participants who exceed the critical z value from among the continuous variables.

Kolmogorov-Smirnov normality test was applied for evaluating whether the data have normal distribution or not since the number of participants was above 50 and skewness levels were examined. The basis for skewness values was taken as the ± 1 interval. The skewness value for some continuous variables was above 1 and it was determined that they do not put forth a normal distribution. Logarithmic transformation was applied to the whole data set for some continuous variables skewed to the right which cannot pass the *Kolmogorov-Smirnov* normality test. It was assumed following the transformation carried out that the skewness values were in the ± 1 interval, meaning that the scores put forth a normal distribution. Thus, it was ensured that the data set meets the required assumptions for the application of ANOVA, MANOVA and *Pearson Product Moments Correlation Analysis* techniques.

According to the ANOVA result, the main effects of age ($F_{(4, 287)}=9.69$, $p<0.001$, $\eta_p^2=0.12$) and education level ($F_{(2, 287)}=198.30$, $p<0.001$, $\eta_p^2=0.58$) variables on the total score of the BNT-60 (TR) are statistically significant. On the contrary, the effect of gender ($F_{(1, 287)}=0.02$, $p>0.05$, $\eta_p^2=0.00$) on the total score of the BNT-60 (TR) was not statistically significant. Hence, the data were combined over the gender variable and the norm values related with the BNT-60 (TR) total score were determined according to the age and education level (Table 2).

According to the 5x2x3 MANOVA results for the BNT-60 (TR) sub-scores (total score, number of spontaneously given correct responses, number of semantic cues given, number of correct responses to semantic cues, number of phonemic cues given, number of correct responses to phonemic cues, number of multiple choice and number of correct responses to multiple choice), the main effects of age (*Pillai's Trace*=0.21, $F_{(4, 287)}=2.24$, $p<0.001$, $\eta_p^2=0.05$) and education level (*Pillai's Trace*=0.61, $F_{(2, 287)}=17.82$, $p<0.001$, $\eta_p^2=0.31$) variables were statistically significant; however the main effects of gender (*Pillai's Trace*=0.02, $F_{(1, 287)}=0.79$, $p>0.05$, $\eta_p^2=0.02$) were not observed to be statistically significant. Therewith, the data were combined over the gender variable and the norm values for

Table 2. Norm Table in Relating to BNT-60 (TR) Total Scores According to Age and Education Level (N= 317)

Education Level	Age Groups				
	60–64 age	65–69 age	70–74 age	75–79 age	80+ age
0–5 year	41.23 (5.20) n=22	40.64 (4.79) n=22	36.59 (3.65) n=22	35.45 (3.22) n=20	31.81 (5.89) n=21
6–11 year	47.85 (4.23) n=26	48.18 (5.50) n=22	43.57 (5.60) n=21	45.25 (4.67) n=20	42.83 (8.77) n=18
12+ year	51.24 (3.25) n=21	49.72 (4.64) n=25	51.00 (5.11) n=23	51.25 (3.58) n=20	49.71 (6.00) n=14

the BNT-60 (TR) sub-scores were determined according to the age and education level (Table 3).

5x2x3 MANOVA was applied in order to determine the probable effects of the age, gender and education level variables on the error scores of the BNT-60 (TR). Only 3 error scores (verbal, multi-word and perceptual) were evaluated as dependent variable with regard to the suitability to parametric analyses. According to the MANOVA results, the main effects of age (*Pillai's Trace*=0.13, $F(4, 287)=3.28$, $p<0.001$, $\eta_p^2=0.04$) and education level (*Pillai's Trace*=0.30, $F(2, 287)=17.15$, $p<0.001$, $\eta_p^2=0.15$) variables were statistically significant; whereas the main effects of the gender (*Pillai's Trace*= 0.01, $F(1, 287)=0.86$, $p>0.05$, $\eta_p^2=0.01$) were not observed to be statistically significant. Therewith, the data were combined over the gender variable and norm values related with the BNT-60 (TR) error scores were determined according to age and education level (Table 4).

When the correlation between the BNT-60 (TR) total score and age is examined, it was observed that there is a negative and weak ($r=-0.25$, $p<0.01$) statistically significance correlation between them.

DISCUSSION

In the BNT-60 (TR), 29 of the original test items were replaced with new items with concrete justifications in accordance with the literature. It is known that similar applications have been made during the adaptation of the BNT to different cultures/languages. For example, 20 items in the Portugal sample group (1), 50 items in the Korean sample group (9), 2 items in the Australia sample group (10), 2 items in the New Zealand sample group (14), 4 items in the Greece sample group (15) have been changed due to cultural/lingual unconformity.

The results acquired from the norm determination study indicate that age and education level have statistically significant effects on the BNT-60 (TR) total scores obtained from the Turkish elderly sample group. However, contrary to expectation, the main effect of gender was not determined as statistically significant, meaning that there is no difference between the female and male participants with regard to the BNT-60 (TR) scores. The reason why gender effect was not statistically significant in the TR version may be due to the fact that items that might cause bias were excluded when forming the BNT-60 (TR) version and that the new items added were selected from among those that had no gender bias. Nonetheless, a statistically significant difference was determined between the BNT-60 (TR) total scores of participants with different age and education levels in accordance with literature. The total score average of the participants in the 60–64 age group was greater in comparison with those of the participants in other age groups, whereas the total mean score of the participants with 12+ years of education, that is those with higher education, was greater in comparison with the participants with different

education levels (Table 2). The decrease in the total score that occurs with aging was in accordance with the norm determination studies carried out in different countries/cultures/languages such as Korea (9), Greece (15), France (16). The strong effect of the education level on the total score of the test is supportive of the previous norm studies (6,12). The results of this study proved that age and education level are important variables that determine the naming ability after the age of 60. The highest total mean score was obtained from the 75–79 age group meaning those who have received higher education of 12+ years ($\bar{X}=51.25$), the lowest total mean score was obtained from the 80+ age group meaning those with low education of 0–5 years ($\bar{X}=31.81$). According to the BNT Korea norm data shows the highest similarity with the methods used when generating the BNT-60 (TR) version, the highest mean score was obtained from the 15–44 age group meaning those who have received 13+ years of education ($\bar{X}=53.93$), whereas the lowest mean score was obtained from the 75+ age group with no education ($\bar{X}=27.42$) (9). Low total scores have been explained by low education level and cultural level in the relevant literature which is important for putting forth the strong effect of education on test scores.

When the means related with the sub-scores of the BNT-60 (TR) are examined, it was determined that participants in the 80+ age group made less use of the phonemic cue provided; whereas they benefited more from the multiple-choice items. Whereas number of spontaneous reactions was high for the participants with an education level of 12+ indicating high education, participants with lower education levels of 0–5 years were less successful in spontaneous naming and needed different cue types (semantic cue, phonemic cue, multiple-choice) more (Table 3). Findings related with the cues are in accordance with the results of Mansur et.al. (13) and Au et.al. (36).

It should be known that the error types made by individuals with cognitive disorder are also important in addition to the total number of correct responses when interpreting the BNT results in clinical use. Hence, the types of erroneous reactions of healthy individuals and patients with cognitive disorders should be determined. Qualitative analyses related with the type and frequency of the errors of healthy individuals in naming enable the differentiation of different patient groups (37,38). According to the BNT-60 (TR) error score analyses, elderly adults with low education, meaning an education ranging between 0–5 years general tend to make more errors in all error types. It was determined that the maximum errors were made in the verbal error category and that this error type decreased with increasing age in the 80+ age group which is in accordance with the results from relevant literature (Table 4).

It is thought that the BNT-60 (TR) will make significant contributions to the objective and reliable evaluation of the naming ability of the elderly population whose native language is Turkish. Required meticulousness

Table 4. Norm Table in Relating to BNT-60 (TR) Error Scores According to Age and Education Level (N= 317)

BNT-TR Error Scores	Education Level	Age Groups				
		60-64 age	65-69 age	70-74 Age	75-79 age	80+ age
Verbal Error	0-5 year	5.95 (2.19) n=22	6.73 (2.51) n=22	6.36 (2.40) n=22	6.70 (2.83) n=20	3.86 (2.35) n=21
	6-11 year	4.77 (1.97) n=26	4.95 (2.79) n=22	5.43 (2.25) n=21	6.20 (2.57) n=20	3.72 (2.24) n=18
	12+ year	4.29 (1.62) n=21	4.24 (2.65) n=25	3.57 (2.48) n=23	3.00 (1.86) n=20	2.50 (2.60) n=14
Multi-word Error		60-64 age	65-69 age	70-74 Age	75-79 age	80+ age
	0-5 year	3.73 (2.21) n=22	2.55 (1.84) n=22	2.55 (1.87) n=22	2.80 (2.24) n=20	2.29 (2.00) n=21
	6-11 year	2.69 (1.98) n=26	2.18 (2.06) n=22	3.52 (1.63) n=21	2.20 (1.61) n=20	2.44 (2.48) n=18
	12+ year	1.71 (1.45) n=21	1.44 (1.53) n=25	1.83 (2.04) n=23	0.95 (1.10) n=20	1.36 (1.50) n=14
Perceptual Error		60-64 age	65-69 age	70-74 Age	75-79 age	80+ age
	0-5 year	2.95 (1.86) n=22	3.91 (2.18) n=22	4.82 (2.70) n=22	4.30 (2.60) n=20	3.95 (1.94) n=21
	6-11 year	2.23 (1.56) n=26	2.55 (2.48) n=22	2.90 (1.70) n=21	2.90 (2.17) n=20	2.17 (1.65) n=18
	12+ year	1.95 (1.86) n=21	1.68 (1.84) n=25	1.61 (1.88) n=23	1.40 (1.73) n=20	1.00 (1.18) n=14

was shown when generating the BNT-60 (TR) version so that the items would be suited to our day, culture and that they included no bias. A sample group that is as homogeneous and clean with regard to subject properties was acquired when collecting norm data. When the elderly sample group of our study was compared with its counterparts, it was observed that it is quite large and qualitatively representative. The elderly sample group of the study was detailed so as to enable performance evaluation in five different age intervals (60-64, 65-69, 70-74, 75-79, 80+ age) taking into consideration the estimated future increase in the elderly population in our country.

As was stated in the introduction section, there is an R&D study that was carried out previously for the BNT for a Turkish sample group between the age range of 20-79 (26). The methodological differences between this aforementioned study and the current study have been listed below as items in a comparative manner: a) The study carried out by Kurt et.al. (26) encompasses the age range between 20-79 with a total of 315 participants. The age range of the examined groups in the elderly sample was 10 (e.g. 60-69 age range). However, the number of participants aged 60 and above in the study was 317. Whereas the age range of the examined groups was kept narrower at 5 (e.g. 60-64 age range). b) Whereas Kurt et.al. (26) used only MMSE for elderly sample group selection, a more detailed screening was made in our study and MOCA, FAQ and GDS were used. c) The determination of the names used for each test item (picture) in the study carried out by Kurt et.al. (26) and the calculation of the difficulty indicators were not carried out and hence, original test starting and ending criteria were not used, however, since our study met these requirements, the original test starting and ending criteria were used in accordance with the original. d) Whereas Kurt et.al. (26) removed only 5 items from the test, the standardization study was carried out for 55 items since they did not suggest new items for the removed items. However, 29 items were removed from the test in our study which were replaced by 29 suggested new items and so, the standardization was carried out for the 60 items in accordance with the original test for each item according to the new item arrangement determined depending on the usage frequencies and difficulty analysis carried out. e) Kurt et.al. (26) kept the usage frequency of the test items

in our language limited with only the data of the participants of the adaptation study (n=57). Whereas in our study, the number of participants who were included in the adaptation study (30 healthy adults: 90 healthy adults between the ages of 19-59: 60-81) was 120 people in total. f) Whereas Kurt et.al. (26) used a different scoring (Spontaneous, Semantic Cue and Vocalic Cue), our study stayed loyal to original scoring principles (Total score: Spontaneous and semantic cue, Sub-score: Semantic, phonemic and multiple-choice cues, error score). g) Since the BNT is a test that evaluates the linguistic ability, test items were determined after an evaluation regarding not only cultural properties, but also regarding 9 critical criteria for linguistic properties that might be critical with regard to Turkish language (the criteria used have been presented with details in the "Method" section). However, there is no such detailed technical evaluation in the study by Kurt et.al. (26) with regard to properties of the Turkish language. h) Whereas there was no information in the study by Kurt et.al. (26) on the adaptation study regarding multiple-choice BNT items, required adaptation and standardization study was carried out in our study for the items with choices.

A comprehensive adaptation of the BNT used widely for determining language disorders in the international and national platforms, norm values were determined for the healthy elderly group whose native language is Turkish (Tables 2, 3 and 4) and was made ready for the use of the related people in our country.

In future studies, the norm values for the age groups other than elderly groups can be determined for the BNT-60 (TR), validity and reliability studies can be carried out for patient groups characterized by linguistic disorders (aphasia, dementia, etc.) and cutoff scores can be determined for the aforementioned diseases.

Ethics Committee Approval: For this study, Hacettepe University Senate Ethics Commission Approval was obtained with decision dated 06.04.2015 and numbered 76000869 / 433-1060.

Informed Consent: Written informed consent was obtained from all participants.

Peer-review: Externally peer-reviewed.

Author Contributions: Concept - BC; Design - AES; Supervision - BC; Resource - AES; Materials - AES; Data Collection and/ or Processing - AES; Analysis and/or Interpretation - AES; Literature Search - AES; Writing - AES, BC; Critical Reviews - BC.

Conflict of Interest: No conflict of interest was declared by the authors.

Financial Disclosure: The authors declared that they did not receive financial support for this study.

Acknowledgment: We thank Prof. Dr. Kaynak Selekler for giving the authors support in providing participants for the research.

REFERENCES

- Miotto EC, Sato J, Lucia M, Camargo CH, Scaff M. Development of an adapted version of the Boston Naming Test for Portuguese speakers. *Rev Bras Psiquiatr* 2010;32:279–282.
- Kimbarow ML, Vangel SJ, Lichtenberg PA. The influence of demographic variables on normal elderly subjects' performance on the Boston Naming Test. *Clin Aphasiol* 1996;24:135–144.
- Tuokko H, Hadjistavrogoulos T. *An Assessment Guide to Geriatric Neuropsychology*. New York: Psychology Press; 1998.
- Lansing AE, Ivnik RJ, Cullum CM, Randolph C. An empirically derived short form of the Boston naming test. *Arch Clin Neuropsychol* 1999;14:481–487.
- 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;16:555–574. [CrossRef]
- Allegri RF, Villavicencio AF, Taragano FE, Rymberg S, Mangone CA, Baumann D. Spanish Boston Naming Test norms. *Clin Neuropsychol* 1997;11:416–420. [CrossRef]
- Ardila A, Rosselli M, Puente AE. *Neuropsychological evaluation of the Spanish speaker*. New York: Plenum Press; 1994. [CrossRef]
- Marien P, Mampaey E, Vervaeke A, Saerens J, De Deyn PP. Normative data for the Boston Naming Test in native Dutch-speaking Belgian elderly. *Brain Lang* 1998;65:447–467. [CrossRef]
- Kim H, Na DL. Normative data on the Korean version of the Boston Naming Test. *J Clin Exp Neuropsychol* 1999;21:127–133. [CrossRef]
- Cruice MN, Worrall LE, Hickson LM. Boston Naming Test results for healthy older Australians: A longitudinal and cross-sectional study. *Aphasiology* 2000;14:143–155. [CrossRef]
- Piguet O, Millar JL, Bennett HP, Lye TC, Creasey H, Broe GA. Boston Naming Test: Normative data for older Australians. *Brain Impair* 2001;2:131–139.
- Tallberg IM. The Boston Naming Test in Swedish: normative data. *Brain Lang* 2005;94:19–31. [CrossRef]
- Mansur LL, Radanovic M, Araújo GDC, Taquemori LY, Greco LL. Boston Naming Test: Performance of Brazilian population from São Paulo. *Pro Fono* 2006;18:13–20. [CrossRef]
- Barker-Collo S. Boston Naming Test performance of older New Zealand adults. *Aphasiology* 2007;21:1171–1180. [CrossRef]
- Patricacou A, Psallida E, Pring T, Dipper L. The Boston Naming Test in Greek: Normative data and the effects of age and education on naming. *Aphasiology* 2007;21:1157–1170. [CrossRef]
- Roberts PM, Doucet N. Performance of French-speaking Quebec adults on the Boston Naming Test. *Can J Speech Lang Pathol* 2011;35:254–267.
- 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;20:828–834. [CrossRef]
- Neils J, Baris JM, Carter C, Dell'aira AL, Nordloh SJ, Weiler E, Weisiger B. Effects of age, education, and living environment on Boston Naming Test performance. *J Speech Hear Res* 1995;38:1143–1149.
- Nebreda MC, García-Caballero A, Asensio E, Revilla P, Rodríguez-Girondo M, Mateos R. A short-form version of the Boston Naming Test for language screening in dementia in a bilingual rural community in Galicia (Spain). *Int Psychogeriatr* 2011;23:435–441. [CrossRef]
- Hawkins KA, Bender S. Norms and the relationship of Boston Naming Test performance to vocabulary and education: A review. *Aphasiology* 2002;16:1143–1153. [CrossRef]
- Budd MA. *Boston Naming Test with Latencies (BNT-L)*. Unpublished dissertation. University of North Texas; 2007. p.168.
- TÜİK. *İstatistiklerle Yaşlılar*. Ankara: Türkiye İstatistik Kurumu Yayınları; 2014.
- Cangöz B. *Sağlıklı bilişsel ve psikolojik yaşlanma*. İçinde: Kutsal YG, editör. *Sağlığımız Önceliğimizdir Yaşlanırken Sağlıklı Kalalım*. Ankara: Hacettepe Üniversitesi Yayınları; 2013. ss.75–82.
- Brouillette RM, Martin CK, Correa JB, Davis AB, Han H, Johnson WD, Foil HC, Hymel A, Keller JN. Memory for names test provides a useful confrontational naming task for aging and continuum of dementia. *J Alzheimers Dis* 2011;23:665–671. [CrossRef]
- Zec RF, Markwell SJ, Burkett NR, Larsen DL. A longitudinal study of confrontation naming in the "normal" elderly. *J Int Neuropsychol Soc* 2005;11:716–726. [CrossRef]
- Kurt M, Can H, Karakaş S. Boston Adlandırma Testi Türk formu için araştırma-geliştirme çalışması. *Yeni Symp* 2016;54:6–14. [CrossRef]
- Nasreddine ZS, Phillips NA, Bedirian V, Charbonneau S, Whitehead V, Collin I, Cummings JL, Chertkow H. The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *J Am Geriatr Soc* 2005;53:695–699. [CrossRef]
- Selekler K, Cangöz B, Uluç S. Montreal bilişsel değerlendirme ölçeği (MOBİD)'nin hafif bilişsel bozukluk ve Alzheimer hastalarını ayırt edebilme gücünün incelenmesi. *Türk Geriatri Derg* 2010;13:166–171.
- Pfeffer RI, Kurosaki TT, Harrah CH, Chance JM, Filos S. Measurement of activities of older adults in community. *J Gerontol* 1982;37:323–329.
- Selekler K, Cangöz B, Karakoç E. İşlevsel Faaliyetler Anketi'nin 50 yaş ve üzeri grupta Türk kültürü için uyarılma ve norm belirleme çalışması. *Türk Nöroloji Dergisi* 2004;10:102–107.
- Yesavage JA, Brink TL, Rose TL, Lum O, Huang V, Adey M, Leirer VO. Development and validation of a geriatric depression screening scale: a preliminary report. *J Psychiatr Res* 1983;7:37–49.
- Ertan T, Eker E. Reliability, validity, and factor structure of the geriatric depression scale in Turkish elderly: are there different factor structures for different cultures? *Int Psychogeriatr* 2000;12:163–172.
- Kaplan E, Goodglass H, Weintraub S. *Boston Naming Test*. Philadelphia: Lea & Febiger; 1983.
- Sözbilir M. Madde analizi ve test geliştirme. <https://olcmevedegerlendirme.files.wordpress.com/2010/09/7-madde-analizi-ve-test-gelistirme.pdf>
- Field A. *Discovering Statistics Using SPSS*. California: Sage Publications; 2009.
- Au R, Joung P, Nicholas M, Obler LK, Kass R, Albert ML. Naming ability across the adult life span. *Aging Neuropsychol Cogn* 1995;2:300–311. [CrossRef]
- Tombaugh TN, Hubiey AM. The 60-item Boston Naming Test: norms for cognitively intact adults aged 25 to 88 years. *J Clin Exp Neuropsychol* 1997;19:922–932. [CrossRef]
- Cangöz B. Yaşlılarda nöropsikolojik (nörokognitif) değerlendirme. *Türkiye Klinikleri: Geriatrik Perspektifte Demans Özel Dergisi* 2015;1:13–19.