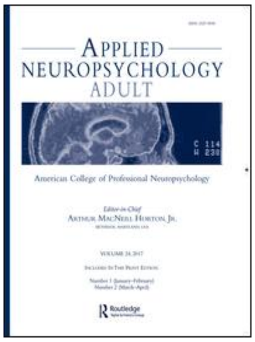

Routledge
Taylor & Francis Group

Applied Neuropsychology: Adult

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/hapn21

Swedish normative data and longitudinal effects of aging for older adults: The Boston Naming Test 30-item and a short version of the Token Test

Katarina Fällman, Ewa Wressle, Jan Marcusson & Elisabet Classon

To cite this article: Katarina Fällman, Ewa Wressle, Jan Marcusson & Elisabet Classon (2022): Swedish normative data and longitudinal effects of aging for older adults: The Boston Naming Test 30-item and a short version of the Token Test, Applied Neuropsychology: Adult, DOI: 10.1080/23279095.2022.2148106

To link to this article: https://doi.org/10.1080/23279095.2022.2148106

© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.

View supplementary material

Published online: 23 Nov 2022.

Submit your article to this journal

Article views: 81

View related articles

View Crossmark data

Full Terms & Conditions of access and use can be found at
https://www.tandfonline.com/action/journalInformation?journalCode=hapn21

APPLIED NEUROPSYCHOLOGY: ADULT
https://doi.org/10.1080/23279095.2022.2148106

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Swedish normative data and longitudinal effects of aging for older adults: The Boston Naming Test 30-item and a short version of the Token Test

Katarina Fällman[a], Ewa Wressle[b], Jan Marcusson[a], and Elisabet Classon[a]

[a]Department of Acute Internal Medicine and Geriatrics and Department of Health, Medicine and Caring Science, Linköping University, Linköping, Sweden; [b]Department of Acute Internal Medicine and Geriatrics and Department of Rehabilitation Medicine in Linköping, Linköping University, Linköping, Sweden

ABSTRACT

Naming ability and verbal comprehension are cognitive functions that may be affected both by normal aging and by disease. Neuropsychological testing is crucial to evaluate changes in language ability and reliable normative data for all ages are needed. We present clinically useful test norms, together with subsample analysis of longitudinal effects of aging, for two robust and well-known tests that evaluate naming ability and verbal comprehension where the present norms for older adults (aged 85 and older) are sparse or missing. Participants ($n = 338$) from a Swedish population-based study, the Elderly in Linköping Screening Assessment, were cognitively evaluated with a cognitive screening battery at the age of 85 years and followed to the age of 93 years. Normative data at age 85 years were calculated from a sample ($n = 207$) that was determined as cognitively healthy after application of rigorous exclusion criteria. Effects of normal aging were investigated by analyzing follow-up performance at age 90 and 93 years for the subsample of cognitively healthy that completed the entire study. The evaluated tests in this study are Swedish versions of the Boston Naming Test 30-item Odd Version (BNT-30) and a short form of the Token Test, Part V (TokV). Analyzes of effects of aging showed that performance decreased with age for BNT-30, but not for TokV. Higher education was associated with better performance in both tests and men performed better than women on the BNT-30. Results also showed naming ability to be more sensitive to aging than verbal comprehension.

KEYWORDS
80 and over; aged; language test; naming ability; neuropsychological tests; verbal comprehension

Introduction

Cognitive functions are known to be affected by normal aging, but with different trajectories. Vocabulary and language abilities in general tend to improve during adulthood up to the sixth or seventh decade. After the seventh decade there is a normal decline in verbal abilities due to aging (Salthouse, 2010) and the decline in e.g., naming abilities seems to be accelerating after the age of 80 (Kent & Luszcz, 2002). However, verbal abilities can also often be affected by disease, such as neurocognitive disorders, leading to communication difficulties for the affected and their relatives. The main risk factor for neurocognitive disorders is high age (Hugo & Ganguli, 2014), and it is important that cognitive tests used to evaluate naming ability and verbal comprehension have valid norms also for older adults, to avoid misjudgments. The forecast is that the population of persons who have reached or surpassed their 8th decade will more than triple globally between 2015 and 2050 (He et al., 2016) and, as the incidence of dementia continues to increase with age (Corrada et al., 2010), valid normative data for persons over 85 years of age is much needed, something that has become a well-recognized current topic (Beker et al., 2019; Kvitting et al., 2019; Melikyan et al., 2019). Many clinical instruments however still lack normative data for older adults and the present study focuses on two assessments measuring naming ability and verbal comprehension for whom this is the case.

The Boston Naming Test (BNT-60; Kaplan et al., 2001; Tallberg, 2005) is one of the most used tests of naming ability world-wide. The BNT-60 and various short forms have been shown to discriminate healthy individuals from persons with neurocognitive disorders (Jefferson et al., 2007; Katsumata et al., 2015) with high positive predictive power (Erdodi et al., 2018) and without practice effects when older individuals (57–85 years of age) are repeatedly tested over three years (Mitrushina & Satz, 1995). The present study evaluates one of several established short versions, the Boston Naming Test 30-item Odd Version (BNT-30; Bezdicek et al., 2021; Fisher et al., 1999; Strauss et al., 2006; Williams et al., 1989). It consists of the 30 items with odd

CONTACT Katarina Fällman katarina.fallman@liu.se Department of Acute Internal Medicine and Geriatrics and Department of Health, Medicine and Caring Science, Linköping University, Medicinska och geriatriska akutkliniken, Universitetssjukhuset, Linköping 581 85, Sweden.

Supplemental data for this article can be accessed online at https://doi.org/10.1080/23279095.2022.2148106

© 2022 The Author(s). Published with license by Taylor & Francis Group, LLC.
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

numbers from the BNT-60 with which it has shown high correlations (Attridge et al., 2022; Fisher et al., 1999; Katsumata et al., 2015) and internal consistencies between different BNT short forms are generally acceptable (Attridge et al., 2022). Swedish normative data for the BNT-30 for older adults are lacking, and normative data have generally been sparse. Preliminary geriatric norms (61–84 years of age) were published in the late nineties but they are based on a small study sample of 15 individuals (Fisher et al., 1999). Bezdicek et al. (2021) recently reported age-norms (up to age 96) for Czech versions of the BNT-30 and BNT-15, finding effect of age, gender, and education. Others have found similar effects in other BNT versions and languages (Vestito et al., 2021; Zec et al., 2007).

The Token Test (Apt, 2008; De Renzi & Vignolo, 1962) is an established assessment of verbal comprehension that includes five subtests with increasing demands on verbal comprehension for every subtest. It has been shown to be a useful indicator of neuropsychological functioning in people with dementia, both as measured by the full test (Taylor, 1998) and by shortened versions (Strauss et al., 2006), it is easy to use and reportedly without practice effects in heathy persons (Lezak et al., 2012). Several short versions are used in clinical practice and have high validity with the original (Strauss et al., 2006). In the current study, a short version used for screening cognitive decline is evaluated. The studied assessment uses six instructions from the original Token Test, Part V (TokV; Nordlund et al., 2011; Stålhammar et al., 2022) and there are no published normative data for older adults for this version.

In the present study, cognitively intact older adults, originating from the Swedish longitudinal population study Elderly in Linköping Screening Assessment (Nägga et al., 2012), were tested repeatedly at age 85, 90 and 93 years. The aim of the study is to present normative values for BNT-30 and TokV for older adults, exploring effects of gender and education, and to evaluate longitudinal age effects on test results for those who participated throughout the whole study. We also investigate whether self-reported vision and hearing impairment affect the results.

Materials and methods

The study participants were originally recruited for the longitudinal population study, Elderly in Linköping Screening Assessment (ELSA 85). In the ELSA 85 study all 85-year-olds (born 1922) living in the municipality of Linköping, Sweden, in 2007 were invited to participate ($n = 650$ persons). The study design is described in Nägga et al. (2012) and the follow-ups in Johansson et al. (2019). Ninety percent ($n = 586$) of the total cohort answered the invitation and 76% ($n = 496$) agreed to participate by answering a postal questionnaire including questions about demographics, education, need for aid, etc. Non-participants were more likely to live in sheltered accommodation or nursing homes, but the health of non-participants in the ELSA 85 study was not greatly different from the participants overall (Dong et al., 2015). Evaluation of cognitive function with neuropsychological assessment (described below) and a somatic checkup by a physician were accepted by 338 persons and were performed during av visit at the Geriatric Clinic at Linköping University Hospital. Follow-up home visits, including repeated neuropsychological assessments, were made when the participants were 90 and 93 years old. The participants were also asked if they experienced any vision or hearing impairment (yes or no) at each visit.

During all phases of the study, the participants could cancel or refuse participation in one or more of the assessments and then continue to another test. The participants also had the possibility to end the examination and continue at another time, if so desired. Only participants that had completed all previous steps of the study were included in each successive follow-up and only those who had fulfilled the Mini Mental State Examination (MMSE; Folstein et al., 1975) were included in this study (Figure 1). At 90 years of age, 45% ($n = 153$) of the participants were deceased, 17% ($n = 58$) declined participation and 4% ($n = 14$) did not answer or were excluded from the study for other reasons (e.g., moved to another municipality). At 93 years of age, an additional 26% ($n = 29$) were deceased and 21% ($n = 24$) declined participation.

The aim of this study is to find normative values for cognitively intact persons. Participants with somatic diseases known to potentially affect cognitive functions such as dementia, mild cognitive impairment, stroke, advanced diabetes or chronic obstructive pulmonary disease were excluded. The exclusion criteria are described in detail in Fällman et al. (2019). Information about the participant's health was obtained from the somatic checkup, from the participants, and from their medical records, collected according to a protocol. When detailed somatic information was not available the participant was excluded.

To further minimize the risk of including persons with cognitive decline, persons with MMSE less than 24 points, i.e., the traditional cutoff (Creavin et al., 2016) were excluded. A higher cutoff point with MMSE less than 26 points has been suggested for persons aged 85 and above (Kvitting et al., 2019) but analyses based on participants with MMSE 26 points or more did not affect the norms in a clinical important way (see Supplementary Material, Tables A and B, for more information) and the cutoff of 24 points was retained (24 points corresponds approximately to the 15th percentile in Swedish norms, depending on age; Kvitting et al., 2019).

The participants' cognitive function was evaluated using the Cognitive Assessment Battery (CAB; Nordlund et al., 2011). CAB is a test battery for screening cognitive functions, developed by the Institute of Neuroscience and Physiology at the University of Gothenburg, Sweden. CAB consists of multiple well-known assessments, some of them as short versions. The tests studied in this paper are the short, odd items version of the Boston Naming Test (BNT-30; Fisher et al., 1999) and a short version of the subtest V from the Token Test (TokV; Nordlund et al., 2011). The assessments were conducted by an occupational therapist, a nurse or a physician from the Geriatric Department of the

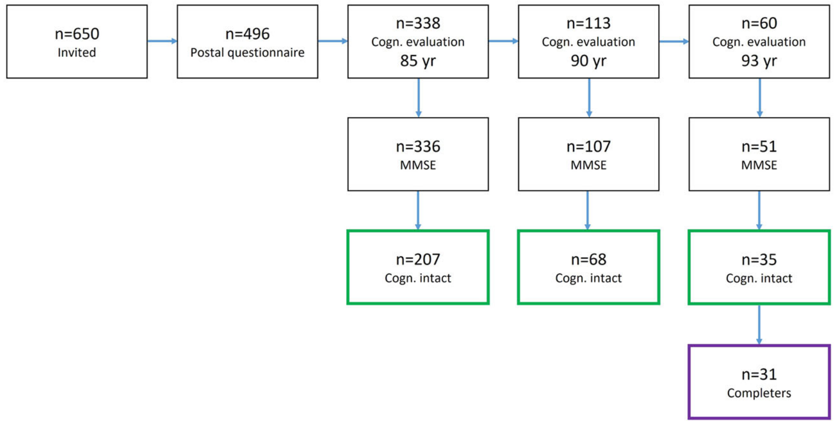

Figure 1. Flowchart of the participants of the Elderly in Linköping Screening Assessment. Cogn. evaluation = participants who were evaluated cognitively. MMSE = participants who completed the Mini Mental State Examination. Cogn. intact = participants who did not meet any exclusion criteria and were further evaluated for the normative data. Completers = Participants with test results from all test sessions.

University Hospital of Linköping. All the investigations were compliant with the instructions in the CAB manual.

The BNT-30 uses the same 30 pictures (items) as the English versions. The full BNT-60 was adapted to Swedish by Tallberg (2005) who used response analysis of the actual lexical items chosen by the norming sample. That study concluded that the original pictures could be retained in the Swedish version. The participant is shown black and white line drawings of objects to be named that represent a range from simple, high frequency vocabulary words to rare words. Administration starts with item 15 and if the participant can correctly name the first four consecutive items (15–18) within the allotted time (20 s per picture), items 1–14 are not administered and automatically awarded one point per item. If the participant cannot correctly name one of items 15–18, the instructor will switch and continue a backward administration, starting with item 14, until the participant has correctly named four items in a row. Automatic points are given for the items not administered and the instructor switches back to continue the forward administration from the item following that at which it had been aborted. Scoring is one point per automatically credited or correctly named item, including items named after stimulus cueing (a clue, for example, “it is something you can eat,” is given when the participant has clearly misperceived the picture). This established administration procedure minimizes demands on the test-taker and closely follows that of the original BNT-60 (Kaplan et al., 2001; Strauss et al., 2006; Stålhammar et al., 2016). Maximum score is 30 with higher score denoting better naming ability.

The TokV stimulus material consists of eight tokens in four different colors (blue, white, red, and green) and two shapes (circle and square). The instructor places the tokens on the table and one by one gives the participants a total of six instructions of varying complexity (e.g., “touch the red circle and the blue square”) to be executed by the participant. The instruction can be repeated once and are the same as instruction numbers 12, 5, 21, 20, 6 and 14 (in that order) from the original Token Test, Part V (Strauss et al., 2006). Outcome measure is number of correctly executed actions, maximum score is 6 points, and higher scores indicate better verbal comprehension.

The ELSA 85 study has been approved by the Research Ethics Committee of Linköping University, Sweden (2006: 141-06, 2012: 332-31; 2014: 455-31) including permission to obtain data from all registers held by the County Council of Östergötland. The study was conducted in accordance with the Declaration of Helsinki on experiments on humans (World Medical Association Declaration of Helsinki, 1997) and written informed consent was collected from all participants at each step of the study.

Mean values, standard deviations, medians, and percentiles were calculated for each age group and for each age in relation to gender and educational level. The 3rd and 16th percentiles are used, corresponding to the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) recommended cutoffs for mild and major neurocognitive disorder (Sachdev et al., 2014). Parametric statistics were chosen due to the large sample size and adherence with previous literature (e.g., Fisher et al., 1999; Jefferson et al., 2007). Shared variance between BNT-30 and TokV was calculated with Pearson’s correlation. Univariate ANOVAs were used to analyze effects of gender, education, and self-reported visual or hearing impairment. Longitudinal age effects were explored by repeated measures using the test results from the subsample of participants with test results from all the three test occasions and fulfilling criteria for cognitive health (referred to below as Completers; $n = 31$, please see Figure 1). Sphericity was calculated using Mauchly’s test and when sphericity could not be assumed the significance was determined using the Greenhouse-Geisser correction. Post hoc analyses between the age groups were carried out using the Bonferroni method (Field, 2018). The level of significance was set at 0.05. All statistical analysis were carried out using SPSS Version 28. If a participant did not complete a test, it is not included in the analysis. Missing data for educational level ($n = 3$) were imputed using the group mean.

Descriptive information, test results as mean values and standard deviations for the whole ELSA 85 population, none excluded, are presented in the Supplementary Material, Table C–F and will not be discussed further in this article.

Table 1. Descriptive information on the participants.

		85-Year-olds *n* (%)	90-Year-olds *n* (%)	93-Year-olds *n* (%)	Completers[a] *n* (%)
Gender	Female	114 (55)	36 (53)	14 (40)	12 (39)
	Male	93 (45)	32 (47)	21 (60)	19 (61)
Education	≤9	147 (71)	49 (72)	24 (69)	20 (65)
(years)	≥10	60 (29)	19 (28)	11(31)	11 (35)
Living situation	Living alone	111 (54)	38 (56)	25 (71)	21 (68)
	Living with family member	96 (46)	30 (44)	10 (29)	10 (32)
Hearing impairment	Yes	139 (67)	52 (76)	30 (86)	28 (90)
Visual impairment	Yes	158 (76)	61 (90)	33 (94)	30 (97)
Total		207	68	35	31

[a]Completers—the participants fulfilling criteria for cognitive health and with test results from all the three test occasions.

Table 2. Normative data for the Boston Naming Test 30-item Odd Version (BNT-30).

		Boston Naming Test 30-item Odd Version (BNT-30)		
Age		85	90	93
Number		206	62	34
Mean		24.4	24.0	24.1
Median		25	25	25
SD[a]		3.9	4.3	3.8
Percentile	3	16	14.5	17
	16	21	19	18.5
	50	25	25	25
	84	28	28	28
	97	30	30	29

[a]SD: standard deviation.

Table 3. Normative data for the subtest of Token Test, Part V (TokV).

		Token Test subtest of Part V (TokV)		
Age		85	90	93
Number		206	59	33
Mean		4.7	5.2	5.0
Median		5	6	5
SD[a]		1.2	1.2	1.2
Percentile	**3**	2	1	2
	16	4	4	3.5
	50	5	6	5
	84	6	6	6
	97	6	6	6

[a]SD: standard deviation.

Results

Descriptive information about the participants is shown in Table 1. All participants had Swedish as their native language and their average school attendance was 8.6 years (SD = 3.6). In initial analyses, education was divided into three levels roughly agreeing with the Swedish school system this age-cohort experienced: less than or equal to 9 years of schooling, 10–12 years, and 13 years and above. However, as there were no significant differences in test results between the two higher educational levels (Bonferroni post hoc test, results not shown), these were combined into one group resulting in two educational levels: less than or equal to 9 years of schooling, or 10 years and above.

Presence of self-reported vision or hearing impairment had no significant impact on either BNT-30 or TokV performance at 85 years of age. Because most of the participants reported impairments at the age of 90 and above, no analyses concerning vision and hearing were performed on the results from the follow-ups.

The correlation between BNT-30 and TokV was $r = .318$, $R^2 = .101$ ($p < .001$). This corresponds to an intermediate effect (Cohen, 1988) showing that the two tests measure shared, but mainly unique, language abilities.

At age 85, men performed better on the BNT-30 than women ($M_{men} = 25.5$, $M_{women} = 24.1$, $F(1, 202) = 5.60$, $p = .019$) and a higher educational level was associated with better performance ($M_{\geq 10} = 25.5$, $M_{\leq 9} = 24.1$, $F(1, 202) = 5.90$, $p = .016$). There was no gender by education interaction. Interaction effects at age 90 and 93 were not explored due to the small number of female participants with 10 years of education or more ($n_{90} = 5$, $n_{93} = 1$). At age 90 there were no significant differences between men and women nor between the two educational groups. At age 93 there was no significant difference between men or women, but higher educational level was again associated with better performance ($M_{\geq 10} = 27.2$, $M_{\leq 9} = 22.9$, $F(1, 30) = 4.35$, $p = .046$). Table 2 displays normative BNT-30 values. For results stratified by education and gender, please see Table G in Supplementary Material.

The results of the TokV were not significantly affected by gender at any age. Higher level of education was associated with better performance at age 85 ($M_{\geq 10} = 5.0$, $M_{\leq 9} = 4.6$, $F(1, 202) = 4.32$, $p = .039$) but there were no significant effects of educational level at age 90 or 93. There was no education by gender interaction at age 85. Normative data for the TokV are shown in Table 3. Please see Table H in Supplementary Material for results stratified by education for TokV at age 85.

There was an age-related decline in test results for BNT-30 for the Completers ($n = 31$) with a mean decrease of two words from the age of 85 to the age of 93 ($M_{85} = 26.5$, $M_{93} = 24.5$, $F(1.50, 44.88) = 10.46$, $p = .001$). Inter individual differences increased with age ($SD_{85} = 2.6$, $SD_{93} = 3.7$) and while 64% ($n = 20$) showed declining performance, 13% ($n = 4$) did not, and 23% ($n = 7$) improved their results. There were no significant age-related changes in TokV test performance ($n = 30$) (Figure 2).

Discussion

We present Swedish normative data for two cognitive assessments measuring naming ability and verbal comprehension derived from a population of cognitively healthy persons 85 years of age and older. The two tests, the Boston Naming Test 30-item Odd version (BNT-30) and a short version of

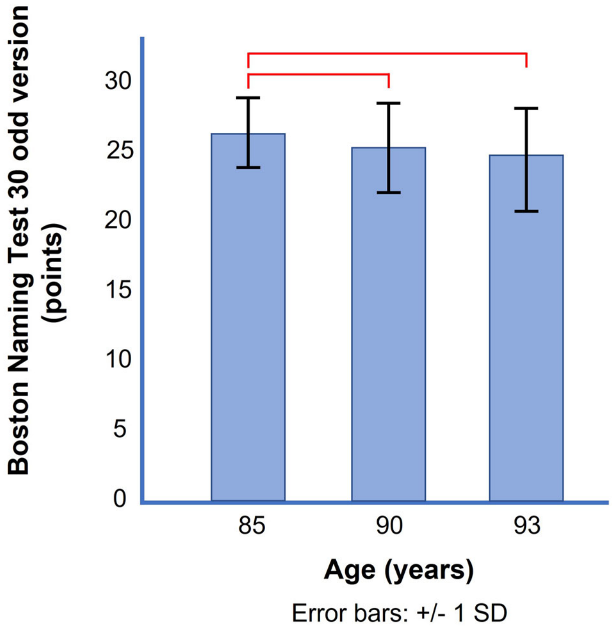

Figure 2. Age effects calculated with repeated measures for the Completers (participants with test results for all three occasions). Error bars with ±1 standard deviation. Lines above the bars indicate significance, p-value < .05.

the Token Test, Part V (TokV), that were found to evaluate different aspects of language. Both are short and easy to administer and regularly used in clinical settings for screening suspected cognitive decline. The current study is one of the first and the largest to present normative data for the population of persons aged 85 years. The presented data is derived from a larger population than the existing clinically used norms and should be generalizable for older persons in countries with similar social contexts and educational systems as in Sweden. The norms apply to the administration method used (starting with item 15; Stålhammar et al., 2016).

BNT-30 mean score at age 85 was 24.4 (SD = 3.9); a lower result and a larger variance both as compared with the previous preliminary norms (M = 27.3, SD = 1.54) that are now more than 20 years old and based on a smaller and younger sample ($n = 15$, age 61–84; Fisher et al., 1999) and compared to the test standards in the manual of the CAB (Nordlund et al., 2011). In CAB the recommended cutoffs are M = 26 and SD = 2 for 80 to 89-year-olds. If following these recommended cutoffs 42% ($n = 87$) of the present cognitively healthy sample would, at age 85, have been classified as outside the norm i.e., performing >1SD below reference mean, 48% ($n = 30$) at age 90 and 41% ($n = 14$) at age 93. Compared to the common clinical approach, to use −1,5 SD as cutoff possible mild cognitive impairment, 33% ($n = 67$) of the study population would be in need of futher evaluation at age 85, 42% ($n = 26$) at age 90 and 35% ($n = 12$) at age 93.

Bezdicek et al. (2021) published normative values for BNT-30 with mean results comparative to the presented norms in this study. The means for the highly educated group of 75–96 years of age in that study are slightly higher, which possibly could be explained by the large age-span. In an American normative study of the even items of BNT, a short version highly correlated to the BNT-30 odd items (Katsumata et al., 2015), Jefferson et al. (2007) also found a higher mean result and smaller variance (M = 28.9, SD = 1.4) for their 85+ year-old participants than those of the present study. But again, their sample was small ($n = 16$) and all but one scored in the high average or better on a measure of literacy and quality of education.

In the present study, as in many normative studies evaluating versions of the Boston Naming Test, higher educational level was associated with better BNT-30 performance (Bezdicek et al., 2021; Jorgensen et al., 2017; Saxton et al., 2000; Tallberg, 2005; Zec et al., 2007). This was however not seen for the 90-year-olds in the current sample, and results from other studies are contradictory regarding education; these are summarized in Kent and Luszcz (2002).

There was also a gender effect such that men outperformed women, which is in line with some (Bezdicek et al., 2021; Jefferson et al., 2007; Zec et al., 2007), but not all previous studies (Saxton et al., 2000; Soylu & Cangoz, 2018; Tallberg, 2005). The effects of education and gender were, in this study, independent which contrasts with the findings of education by gender interactions in Jefferson et al. (2007).

In the cross-sectional studies by Jorgensen et al. (2017) Soylu and Cangoz (2018) and Bezdicek et al. (2021) there was an age-related decrease in test results with BNT-60 or BNT-30, whereas Tallberg (2005) and Jefferson et al. (2007) found no significant effect of age. Two longitudinal studies by Zec et al. (2005) and Kent and Luszcz (2002) showed decreasing performance with advancing age at repeated testing over ten to eight years using BNT-60 and short versions of 15 items, well in line with the age-related decrease in performance found in the present study.

The TokV version of the Token Test studied in this article is used in clinical practice and research mainly as a part of the CAB (Nordlund et al., 2011; Stålhammar et al., 2022). The test is short and easy to use which makes it desirable in clinical practice. The present findings also indicate it is relatively robust; there were no effects of gender or age in our sample, but higher education was associated with better performance. To the authors knowledge there are no published normative data for TokV, but the CAB manual states that persons of all ages should manage to execute all six instructions without any mistakes. Following this recommendation would classify 69% of the cognitively intact 85-year-old participants in this study as outside the norm and in need of further investigations. In Fällman et al. (2019) the same population had test results within the existing norms (when comparisons could be made) for other cognitive tests. The discrepancy between the normative values in this study and the norms in the CAB manual, for both BNT-30 and TokV, highlights the importance of age-normative values for all ages because of the risk of misjudgments.

Performance was not significantly affected by self-reported vision or hearing impairment, neither in the BNT-30 or the TokV.

There are some limitations to the study that needs to be addressed. The definition of "cognitively intact" participants was made after the collection of data, resulting in the risk of excluding cognitively healthy participants, for example with somatic disease, as well as the risk of including persons with preclinical cognitive disorder. We did not have information concerning biomarkers such as e.g., beta-amyloid and tau proteins from the participants, but we believe the careful exclusion of participants with conditions known to impact cognition helps ensure that the norms presented reflect healthy aging. However, there is always a risk of biased test results in longitudinal studies of older adults. This study design tends to result in a more homogenous group over the years because the healthier participants are more likely to continue to participate throughout the study. Therefore, the age effect was studied in the subgroup that participated at all time points, reducing the possibility of differences in population between time points.

Ideal normative data are based on large sample sizes to reduce the risk of sampling errors. This is often difficult in older adults and normative data for the nonagenarians in this study are unfortunately based on small sample sizes. Larger normative studies need to be done to confirm our results. However, normative data based on a demographically fit sample may by more representative than data based on a large age heterogeneous sample for a certain patient, and therefore the norms presented in this study are nonetheless useful in clinical practice.

In summary, we present valid and representative normative values for very old persons, in Sweden and other countries with similar social contexts and educational systems, for two language tests: BNT-30 and TokV. We explored the influence of education, gender and self-reported vision and hearing impairment and found that higher education was associated with better performance in both tests and men performed better than women on the BNT-30 at age 85. There was an age-related decline in test results for BNT-30, but not TokV. Self-reported hearing or vision impairment did not impact the results. This is the largest study to date to present norms for these tests for persons aged 85 and older. It highlights the importance of updated normative values for tests used in clinical practice for all ages.

Acknowledgements

We would like to thank the participants for devoting their time and energy to this project. We would also like to thank our colleagues who have taken part in the data collection with enthusiasm and endurance.

Disclosure statement

The authors report there are no competing interests to declare.

Funding

This work was supported by the ALF Grant, Region Östergötland, under Grant [numbers RÖ-790811 and RÖ-900621]. The funders had no input in or influence on the study.

ORCID

Katarina Fällman http://orcid.org/0000-0002-5289-8831
Jan Marcusson http://orcid.org/0000-0002-6452-3930
Elisabet Classon http://orcid.org/0000-0003-1403-6357

References

Apt, P. (2008). *Utprövning av tre svenska versioner av Token Test på en vuxen normalgrupp* [Master thesis]. Lund University. http://lup.lub.lu.se/student-papers/record/2863318

Attridge, J., Zimmerman, D., Rolin, S., & Davis, J. J. (2022). Comparing Boston Naming Test short forms in a rehabilitation sample. *Applied Neuropsychology: Adult*, *29*(4), 810–815. https://doi.org/10.1080/23279095.2020.1811984

Beker, N., Sikkes, S. A. M., Hulsman, M., Schmand, B., Scheltens, P., & Holstege, H. (2019). Neuropsychological test performance of cognitively healthy centenarians: Normative data from the Dutch 100-plus study. *Journal of the American Geriatrics Society*, *67*(4), 759–767. https://doi.org/10.1111/jgs.15729

Bezdicek, O., Rosicka, A. M., Mana, J., Libon, D. J., Kopecek, M., & Georgi, H. (2021). The 30-item and 15-item Boston Naming Test Czech version: Item response analysis and normative values for healthy older adults. *Journal of Clinical and Experimental Neuropsychology*, *43*(9), 890–905. https://doi.org/10.1080/13803395.2022.2029360

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Corrada, M. M., Brookmeyer, R., Paganini-Hill, A., Berlau, D., & Kawas, C. H. (2010). Dementia incidence continues to increase with age in the oldest old the 90+ study. *Annals of Neurology*, *67*(1), 114–121. https://doi.org/10.1002/ana.21915

Creavin, S. T., Wisniewski, S., Noel-Storr, A. H., Trevelyan, C. M., Hampton, T., Rayment, D., Thom, V. M., Nash, K. J., Elhamoui, H., Milligan, R., Patel, A. S., Tsivos, D. V., Wing, T., Phillips, E., Kellman, S. M., Shackleton, H. L., Singleton, G. F., Neale, B. E., Watton, M. E., & Cullum, S. (2016). Mini-Mental State Examination (MMSE) for the detection of dementia in clinically unevaluated people aged 65 and over in community and primary care populations. *Cochrane Database of Systematic Review*, *2016*(1):CD011145. https://doi.org/10.1002/14651858.CD011145.pub2

De Renzi, E., & Vignolo, L. A. (1962). The Token Test: A sensitive test to detect receptive disturbances in aphasics. *Brain*, *85*, 665–678. https://doi.org/10.1093/brain/85.4.665

Dong, H. J., Wressle, E., & Marcusson, J. (2015). Unaltered image of health maintenance: An observation of non-participants in a Swedish cohort study of 85 to 86 year olds. *The Journal of Frailty & Aging*, *4*(2), 93–99. https://doi.org/10.14283/jfa.2015.47

Erdodi, L. A., Dunn, A. G., Seke, K. R., Charron, C., McDermott, A., Enache, A., Maytham, C., & Hurtubise, J. L. (2018). The Boston Naming Test as a measure of performance validity. *Psychological Injury and Law*, *11*(1), 1–8. https://doi.org/10.1007/s12207-017-9309-3

Fällman, K., Lundgren, L., Wressle, E., Marcusson, J., & Classon, E. (2019). Normative data for the oldest old: Trail Making Test A, Symbol Digit Modalities Test. Victoria Stroop Test and Parallel Serial Mental Operations. *Aging, Neuropsycholgy, and Cognition*, *27*(4), 567–580. https://doi.org/10.1080/13825585.2019.1648747

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). Sage Publications Inc.

Fisher, N. J., Tierney, M. C., Snow, W. G., & Szalai, J. P. (1999). Odd/even short forms of the Boston Naming Test: Preliminary geriatric norms. *The Clinical Neuropsychologist*, *13*(3), 359–364. https://doi.org/10.1076/clin.13.3.359.1742

Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). "Mini-mental state": A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, *12*(3), 189–198. https://doi.org/10.1016/0022-3956(75)90026-6

He, W., Goodkind, D., & Kowal, P. (2016). *An aging world 2015*. U.S. Goverment Publishing Office.

Hugo, J., & Ganguli, M. (2014). Dementia and cognitive impairment: Epidemiology, diagnosis, and treatment. *Clinics in Geriatric Medicine*, *30*(3), 421–442. https://doi.org/10.1016/j.cger.2014.04.001

Jefferson, A. L., Wong, S., Gracer, T. S., Ozonoff, A., Green, R. C., & Stern, R. A. (2007). Geriatric performance on an abbreviated version of the Boston Naming Test. *Applied Neuropsychology*, *14*(3), 215–223. https://doi.org/10.1080/09084280701509166

Johansson, M. M., Marcusson, J., & Wressle, E. (2019). Maintaining health-related quality of life from 85 to 93 years of age despite decreased functional ability. *British Journal of Occupational Therapy*, *82*(6), 348–356. https://doi.org/10.1177/0308022619830261

Jorgensen, K., Johannsen, P., & Vogel, A. (2017). A Danish adaptation of the Boston Naming Test: Preliminary norms for older adults and validity in mild Alzheimer's disease. *The Clinical Neuropsychologist*, *31*(sup1), 72–87. https://doi.org/10.1080/13854046.2017.1371337

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *The Boston Naming Test*. PRO-ED.

Katsumata, Y., Mathews, M., Abner, E. L., Jicha, G. A., Caban-Holt, A., Smith, C. D., Nelson, P. T., Kryscio, R. J., Schmitt, F. A., & Fardo, D. W. (2015). Assessing the discriminant ability, reliability, and comparability of multiple short forms of the Boston Naming Test in an Alzheimer's disease center cohort. *Dementia and Geriatric Cognitive Disorders*, *39*(3–4), 215–227. https://doi.org/10.1159/000370108

Kent, P. S., & Luszcz, M. A. (2002). A review of the Boston Naming Test and multiple-occasion normative data for older adults on 15-item versions. *The Clinical Neuropsychologist*, *16*(4), 555–574. https://doi.org/10.1076/clin.16.4.555.13916

Kvitting, A. S., Fällman, K., Wressle, E., & Marcusson, J. (2019). Age-normative MMSE data for older persons aged 85 to 93 in a longitudinal Swedish cohort. *Journal of the American Geriatrics Society*, *67*(3), 534–538. https://doi.org/10.1111/jgs.15694

Lezak, M., Howieson, D. B., Bigler, E. D., & Tranel, D. (2012). *Neuropsychological assessment* (5th ed.). Oxford University Press.

Melikyan, Z. A., Corrada, M. M., Dick, M. B., Whittle, C., Paganini-Hill, A., & Kawas, C. H. (2019). Neuropsychological Test Norms in cognitively intact oldest-old. *Journal of the International Neuropsychological Society*, *25*(5), 530–545. https://doi.org/10.1017/S1355617719000122

Mitrushina, M., & Satz, P. (1995). Repeated testing of normal elderly with the Boston Naming Test. *Aging*, *7*(2), 123–127. https://doi.org/10.1007/BF03324301

Nägga, K., Dong, H. J., Marcusson, J., Skoglund, S. O., & Wressle, E. (2012). Health-related factors associated with hospitalization for old people: Comparisons of elderly aged 85 in a population cohort study. *Archives of Gerontology and Geriatrics*, *54*(2), 391–397. https://doi.org/10.1016/j.archger.2011.04.023

Nordlund, A., Pahlsson, L., Holmberg, C., Lind, K., & Wallin, A. (2011). The Cognitive Assessment Battery (CAB): A rapid test of cognitive domains. *International Psychogeriatrics*, *23*(7), 1144–1151. https://doi.org/10.1017/S1041610210002334

Sachdev, P. S., Blacker, D., Blazer, D. G., Ganguli, M., Jeste, D. V., Paulsen, J. S., & Petersen, R. C. (2014). Classifying neurocognitive disorders: The DSM-5 approach. *Nature Reviews. Neurology*, *10*(11), 634–642. https://doi.org/10.1038/nrneurol.2014.181

Salthouse, T. A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society*, *16*(5), 754–760. https://doi.org/10.1017/S1355617710000706

Saxton, J., Ratcliff, G., Munro, C. A., Coffey, E. C., Becker, J. T., Fried, L., & Kuller, L. (2000). Normative data on the Boston Naming Test and two equivalent 30-item short forms. *The Clinical Neuropsychologist*, *14*(4), 526–534. https://doi.org/10.1076/clin.14.4.526.7204

Soylu, A. E., & Cangoz, B. (2018). Adaptation and norm determination study of the Boston Naming Test for healthy Turkish elderly. *Noro Psikiyatri Arsivi*, *55*(4), 341–348. https://doi.org/10.5152/npa.2017.19331

Stålhammar, J., Hellström, P., Eckerström, C., & Wallin, A. (2022). Neuropsychological test performance among native and non-native Swedes: Second languauge effects. *Archives of Clinical Neuropsychology*, *37*(4), 826–838. https://doi.org/10.1093/arclin/acaa043

Stålhammar, J., Ryden, I., Nordlund, A., & Wallin, A. (2016). Boston Naming Test automatic credits inflate scores of nonaphasic mild dementia patients. *Journal of Clinical and Experimental Neuropsychology*, *38*(4), 381–392. https://doi.org/10.1080/13803395.2015.1119254

Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary*. Oxford University Press.

Tallberg, I. M. (2005). The Boston Naming Test in Swedish: Normative data. *Brain and Language*, *94*(1), 19–31. https://doi.org/10.1016/j.bandl.2004.11.004

Taylor, R. (1998). Indices of neuropsychological functioning and decline over time in dementia. *Archives of Gerontology and Geriatrics*, *27*(2), 165–170. https://doi.org/10.1016/S0167-4943(98)00111-3

Vestito, L., Mori, L., Trompetto, C., Tomatis, M., Alessandria, G., De Carli, F., Cocito, L., & Bandini, F. (2021). The 15-item version of the Boston Naming Test in Italian: Normative data for adults. *Aphasiology*. https://doi.org/10.1080/02687038.2021.1988506

Williams, B. W., Mack, W., & Henderson, V. W. (1989). Boston Naming Test in Alzheimer's disease. *Neuropsychologia*, *27*(8), 1073–1079. https://doi.org/10.1016/0028-3932(89)90186-3

World Medical Association Declaration of Helsinki. (1997). Recommendations guiding physicians in biomedical research involving human subjects. *JAMA*, *277*(11), 925–926.

Zec, R. F., Burkett, N. R., Markwell, S. J., & Larsen, D. L. (2007). Normative data stratified for age, education, and gender on the Boston Naming Test. *The Clinical Neuropsychologist*, *21*(4), 617–637. https://doi.org/10.1080/13854040701339356

Zec, R. F., Markwell, S. J., Burkett, N. R., & Larsen, D. L. (2005). A longitudinal study of confrontation naming in the "normal" elderly. *Journal of the International Neuropsychological Society*, *11*(6), 716–726. https://doi.org/10.1017/S1355617705050897