





The Boston Naming Test-South African Short Form, Part I: Psychometric properties in a group of healthy English-speaking university students

**Authors:**

Kevin G.F. Thomas¹ 
 Lauren Baerecke¹ 
 Chen Y. Pan¹ 
 Helen L. Ferrett² 

Affiliations:

¹Department of Psychology,
 Faculty of Humanities,
 University of Cape Town,
 Cape Town, South Africa

²Department of Psychiatry,
 Faculty of Health Sciences,
 Stellenbosch University,
 Stellenbosch, South Africa

Corresponding author:

Kevin Thomas,
 kevin.thomas@uct.ac.za

Dates:

Received: 14 June 2019
 Accepted: 18 Sept. 2019
 Published: 22 Nov. 2019

How to cite this article:

Thomas, K.G.F., Baerecke, L.,
 Pan, C.Y., & Ferrett, H.L. (2019).
 The Boston Naming Test-South
 African Short Form, Part I:
 Psychometric properties in a
 group of healthy English-
 speaking university students.
*African Journal of Psychological
 Assessment*, 1(0), a15. [https://
 doi.org/10.4102/ajopa.v1i0.15](https://doi.org/10.4102/ajopa.v1i0.15)

Copyright:

© 2019. The Authors.
 Licensee: AOSIS. This work
 is licensed under the
 Creative Commons
 Attribution License.

Read online:

Scan this QR
 code with your
 smart phone or
 mobile device
 to read online.

The Boston Naming Test (BNT) is a popular cognitive test designed to detect word-finding difficulties in neurologic disease. However, numerous studies have demonstrated the BNT's inherent cultural bias and cautioned against uncritical administration outside of North America. There is little research on the BNT performance of South African samples and on ways to make the test culturally fair for use in this country. In this article, we describe the development and psychometric properties of the BNT-South African Short Form (BNT-SASF). This instrument includes 15 items drawn from the original test pool and judged by a panel of practising neuropsychologists and community members to be culturally appropriate for use in South Africa. We administered the standard 60-item BNT and the BNT-SASF to a homogeneous (English-fluent, high socioeconomic status and highly educated) sample of young South African adults. This design allowed us to avoid potentially confounding sociodemographic influences in our evaluation of the instrument's basic utility. We found that the BNT-SASF demonstrates fundamental psychometric properties that are the equivalent of short forms developed elsewhere. Moreover, it appears to measure the same construct as the 60-item BNT while being less culturally biased. We conclude that the BNT-SASF has potential utility in South African assessment settings. It is quick and easy to administer, thus aiding in the rapid screening of patients. Moreover, it is cost-effective because its items are drawn from the pool comprising the original test. Future research will describe psychometric properties of Afrikaans and isiXhosa versions of the BNT-SASF and investigate diagnostic validity in dementia patients.

Keywords: Boston Naming Test; cross-cultural neuropsychology; cultural bias; reliability; short form; validity.

Introduction

The Boston Naming Test (BNT; Kaplan, Goodglass, & Weintraub, 1978, 1983, 2001) is a widely used cognitive test designed to detect the serious word-finding difficulties that characterise certain variants of aphasia and dementia. However, numerous studies have suggested that the BNT is culturally biased and cautioned against uncritical administration of the instrument (Barker-Collo, 2001; Fernández & Abe, 2018). To date, there is little published research on the BNT performance of South African samples and on ways to make the test culturally fair for use in this country.

Boston Naming Test: A brief introduction

The BNT tests confrontation naming ability (i.e. the ability to pull out the correct word at will; Lezak, Howieson, & Loring, 2004, p. 511). In its current form, it consists of 60 black-and-white line drawings presented in ascending order of difficulty. The first few items are commonly encountered objects (e.g. bed), whereas the last several are less frequently encountered objects (e.g. protractor). For each item, the examinee is given 20 seconds to produce a correct spontaneous response, after which a semantic cue is offered (e.g. *it measures angles*). Failing the production of a correct response to this cue, a phonemic cue is offered (e.g. *it starts with the sound 'pro'*). The most recent revision also features a multiple-choice section. After completing the standard presentation as described above, the examiner returns to each failed item and asks the examinee to select, from an array of four options, the word best describing the pictured object.

The BNT is used primarily to assess confrontation naming ability in patients of all ages with neurological deficits stemming from cerebrovascular accidents, traumatic brain injuries and

neurodegenerative disorders (Kiran et al., 2018; Strain et al., 2017; Strauss, Sherman, & Spreen, 2006). It is particularly effective in detecting the naming deficits present in Alzheimer's disease (AD) and thus helps distinguish that neurodegenerative disorder from normal aging and from other forms of dementia (Balthazar, Cendes, & Damasceno, 2008; Golden et al., 2005).

Interpretation of BNT performance is complicated by the fact that non-organic factors may impact on scores. For instance, both age and education moderate BNT performance in healthy individuals. Scores decline with increasing age, with especially significant deterioration in the oldest old (by conventional definition, those aged 80 years and older; Lucas et al., 2005; Tombaugh & Hubley, 1997; Zec, Burkett, Markwell, & Larsen, 2007). Scores are also lower in those with fewer years of education, with particularly strong effects at < 12 years (Hawkins & Bender, 2002; Mitrushina, Boone, Razani, & D'Elia, 2005; Neils et al., 1995).

Cross-cultural adaptation and use of the Boston Naming Test

As is the case with many other popular neuropsychological tests, the BNT was developed for the assessment of monolingual English-speaking North American individuals and reflects the context in which it was developed. Unsurprisingly, then, BNT performance of non-North American samples is markedly poorer than that of North American samples (see, e.g., Cruice, Worrall, & Hickson, 2000; Tallberg, 2005). Perhaps more surprising is that this cross-cultural difference exists even when evaluating the performance of English speakers from New Zealand or Australia against North American normative data, or when comparing the performance of White Americans to that of African-Americans, bilingual Spanish/English residents of the United States, or bilingual French/English residents of Canada (Barker-Collo, 2001; Fillenbaum, Huber, & Taussig, 1997; Kohnert, Hernandez, & Bates, 1998; Lichtenberg, Ross, & Christensen, 1994; Roberts, Garcia, Desrochers, & Hernandez, 2002).

Often, the source of these performance differences is the cultural relevance of items to test-takers. Evidence supporting this statement emerges from studies showing that examinees with different ethnic or cultural backgrounds produce different patterns of errors (Allegrì et al., 1997; Pedraza et al., 2009). Moreover, particular items (e.g. *beaver*, *pretzel*) appear to be especially culturally loaded: in non-North American samples, error rates on those items are significantly higher than those on adjacent items (i.e. items that should have a similar level of difficulty; Barker-Collo, 2007; Worrall, Yiu, Hickson, & Barnett, 1995). Hence, researchers and clinicians across the world have developed culturally modified versions of the test, replacing problematic items with ones more suited to their local contexts (see, e.g., Fernández & Fulbright, 2015; Grima & Franklin, 2016; Kim & Na, 1999; Patricacou, Psallida, Pring, & Dipper, 2007).

The current study

We describe the development of, and present preliminary psychometric data for, the Boston Naming Test-South African Short Form (BNT-SASF). We chose to develop a short (15-item) form because such instruments aid in the rapid screening of patients. Cognitive screening instruments are especially important in the resource-limited and patient-heavy clinics that characterise the South African healthcare system (Katzef, Henry, Gouse, Robbins, & Thomas, 2019; Robbins et al., 2013). Moreover, reduced test time facilitates the assessment of patients with limited attention or motivation, and of those with severe neurological impairment who may become easily fatigued or frustrated (Roebuck-Spencer et al., 2017).

There is an extensive precedent for creating a short form of the BNT (Fastenau, Denburg, & Mauer, 1998; Kang, Kim, & Na, 2000; Saxton et al., 2000). Certain 15-item and 30-item short forms appear to have clinical utility, showing high rates of agreement with the full 60-item test in distinguishing dementia patients from healthy older adults (Graves, Bezeau, Fogarty, & Blair, 2004; Lansing, Ivnik, Cullum, & Randolph, 1999; Williams, Mack, & Henderson, 1989). Additionally, age, education and culture moderate performance on these short forms in the same way they do on the full test (Jefferson et al., 2007; Kent & Luszcz, 2002; Leite, Miotto, Nitrini, & Yassuda, 2017).

We modelled procedures for our short form development on those described by Mack, Freed, Williams and Henderson (1992). They created four equivalent 15-item versions by dividing the 60 items of the original test into four 15-item groups, with each group reflecting the original's full range of content. They reported that each short form successfully differentiated a sample of AD patients from healthy controls. Their fourth version, the Mack SF-4, is the most globally popular 15-item short form and it is included with the officially published BNT kit.

The BNT-SASF comprises 15 items judged by a forum of practising neuropsychologists and community members as being more culturally appropriate for the South African population than those on the Mack SF-4. This article is the first to provide a detailed psychometric report on a version of the BNT designed specifically for use in South Africa. Although Mosdell, Balchin, and Ameen (2010) describe a South African-adapted 30-item form of the BNT, they do not (1) provide reliability or validity information, (2) compare performance on their short form to performance on the full version of the instrument or to performance on previously published short forms, or (3) present item-level analyses. Moreover, their adapted instrument features entirely new items, not included on the original BNT, making it somewhat less accessible to clinicians than the BNT-SASF.

Using a relatively homogeneous sample to minimise the influence on BNT performance of potentially confounding

factors such as age, education and language background, the current study addressed these specific questions:

- How does the BNT performance of English-fluent university undergraduate students compare with North American normative standards?
- Do basic psychometric properties of the BNT, as established in its development literature, hold in this South African sample?
- What is the test–retest and internal consistency reliability of the BNT-SASF?
- Do the items included in the BNT-SASF show the desired properties in terms of, for instance, relative difficulty?

Methods

Development of the Boston Naming Test-South African Short Form

The BNT-SASF comprises 15 items drawn from the BNT's pool of 60 items (Table 1).

To decide which 15 items would constitute the instrument, we consulted via email with 15 fully trained and experienced South African neuropsychologists personally known to us (ten based in the Western Cape, three in Gauteng, one in the Eastern Cape and one in KwaZulu-Natal). All were members of the South African Clinical Neuropsychological Association (SACNA), and all had used the BNT in their clinical practice for several years. We told them we had divided the pool of 60 items into 15 sets of four items of equivalent difficulty (e.g. items 1–4 formed a set, items 5–8 formed another set, and so on; this procedure ensured the items in the short form would be of increasing difficulty and in a sequence roughly equivalent to the original test). We instructed the neuropsychologists to rate each item in each of the 15 sets according to whether it was culturally appropriate for use in South Africa, and to then select the most culturally appropriate of the four items in each set. For instance, the item *beaver* was one of the options in the eighth set. However, this animal is likely to be relatively unfamiliar to the average South African; *rhinoceros* (another option in the same set) is likely to be more culturally appropriate. After taking the

consensus of views, we settled on the final version of the BNT-SASF. A team of linguists translated and back-translated this modified test from English into Afrikaans and isiXhosa, the other two languages most widely spoken in the Western Cape. To ensure that the isiXhosa version was appropriate for use in that province, we consulted with a small forum of community members (five women, aged from the mid-20s to mid-60s, all first-language isiXhosa speakers) from Khayelitsha and Gugulethu. We report in more detail on those versions of the BNT-SASF in forthcoming publications

Participants

We used convenience sampling to recruit and screen 104 undergraduate students. Forty-five did not meet the eligibility criteria listed below. Hence, the final sample consisted of 59 participants (24 men and 35 women). They received course credit in exchange for participation.

Participants were required to (1) be aged between 18 and 25 years; (2) speak English as a first language; (3) have matriculated from a South African Quintile 4 or Quintile 5 public high school (or the relative equivalent if schooled elsewhere)¹ or from a private high school in South Africa, and have gained entry into university; (4) have their home residence in a suburb with a median annual income of ≥ R76 801 (Statistics South Africa, 2011) and (5) make themselves available for one of the research slots listed on the online schedule distributed to them. We set inclusion criteria related to quality of education and socioeconomic status (SES) in place because, although there is not a large literature detailing their influence on BNT performance, numerous studies describe their general and significant relations to cognitive performance (see, e.g., Crowe et al., 2012; Lyu & Burr, 2016).

We excluded individuals with a current prescription for psychotropic medication and/or a history of psychiatric diagnosis; a history of pre-natal or birth complications; a history of head injury that resulted in a loss of consciousness for more than 5 min; seizure disorders; substance-use disorders; a history of medical illness that resulted in loss of cognitive functioning; or language, speech or behavioural disorders. We also excluded those who had been administered psychometric tests in the 12 months prior to study enrolment. Again, we set these exclusion criteria in place because these factors influence cognitive test performance (Mitrushina et al., 2005; Strauss et al., 2006).

Measures and procedure

Each participant was tested individually, across two sessions separated by exactly 2 weeks, in a quiet testing room within

TABLE 1: Items comprising the Boston Naming Test-South African Short Form.

BNT-SASF item number	Object pictured	BNT (full version) item number
1	Tree	2
2	Comb	7
3	Toothbrush	10
4	Hanger	15
5	Bench	20
6	Snail	22
7	Dart	25
8	Rhinoceros	31
9	Dominoes	35
10	Escalator	37
11	Stethoscope	42
12	Funnel	46
13	Compass	50
14	Sphinx	55
15	Protractor	59

BNT-SASF, Boston Naming Test-South African Short Form.

¹Section 35(1) of the South African Schools Act requires that each province's executive council consults each year with the National Minister of Education to identify and publish the national quintile category within which each public school in the province will be placed. A school's quintile is determined by the wealth of the surrounding community (i.e. the likely wealth of most students who will attend the school). Quintile 1 schools are the poorest 20% of schools, Quintile 2 schools are the next poorest 20%, and so on, with Quintile 5 including the wealthiest 20% of schools. Quintile 1 schools receive the highest per-student governmental allocation, and Quintile 5 the lowest. Quintiles 1–3 include no-fee schools (<http://section27.org.za/wp-content/uploads/2017/02/Chapter-7.pdf>).

a psychology research laboratory. A psychology graduate student administered all study procedures.

Test occasion 1 (T1)

Upon entering the laboratory, the researcher ensured the participant read, understood and signed an informed consent document. Before administering the psychological tests, the researcher ensured that the participant completed a study-specific sociodemographic questionnaire. This instrument gathered biographical, socioeconomic and medical information needed for screening purposes.

Those meeting the eligibility criteria were administered the 60-item BNT according to the standardised instructions that appear in the test manual (Kaplan et al., 2001), with this exception: the test administrator presented all 60 items, in order from Item 1 through Item 60 (i.e. the usual starting point and discontinuation rules were not applied). We followed this procedure to ensure that performance on all 60 items could be examined statistically.

The BNT-SASF was not administered as a separate measure to participants. Instead, we derived a score for the instrument from the performance on relevant items within the full BNT administration.

At the end of the test administration, the researcher scheduled an appointment for the second test session.

Test occasion 2 (T2)

Immediately after entering the laboratory, participants were reminded of their research rights and they were then administered the BNT (including, of course, the 15 items that constituted the BNT-SASF).

Ethical considerations

The study protocol was approved by our institution's review board. All procedures were conducted in compliance with the Declaration of Helsinki (World Medical Association, 2013). Our consent document gave participants complete information about the study procedures, assured them of their rights to privacy and to confidentiality of their data and informed them that they could withdraw from the study at any point without penalty. The document also informed them about their course credit compensation and about the minimal risks they would face during participation. Finally, participants were fully debriefed at the end of T2 and given the opportunity to ask any questions relating to their experience of the research.

All study procedures were approved by the University of Cape Town's Department of Psychology Research Ethics Committee (clearance number PSY2019-005).

Data management and statistical analyses

We scored the 60-item BNT and the BNT-SASF using conventional methods (i.e. the total score for each instrument

is the sum of the number of correct spontaneous responses and the number of correct responses following a stimulus cue). We entered those outcome variables, along with the score for each item (0 or 1), into a datasheet. We analysed the data using SPSS (version 25.0), with the threshold for statistical significance (α) set at 0.05.

Analyses of the BNT and BNT-SASF data proceeded across four discrete steps. First, two separate one-sample *t*-tests compared BNT performance of the current sample at T1 to average BNT performance of highly educated young adults from North America and New Zealand; and three separate paired-sample *t*-tests compared the T1 performance of the current sample on the 15 items comprising the BNT-SASF to their T1 performance on 15 items comprising previously established short forms. Second, Spearman's ρ estimated test-retest reliability for each instrument was established across the 2-week interval between T1 and T2. (We used this coefficient, rather than Pearson's product-moment correlation coefficient, because test scores were non-normally distributed.) Third, Cronbach's α estimated internal consistency reliability for the T1 data. Fourth, we investigated item-by-item performance on both instruments by creating a difficulty index for each item (i.e. calculating, for each item across the entire sample, the proportion of correct responses produced either spontaneously or following the presentation of a semantic cue). Several previous BNT studies have calculated the difficulty index in this way (see, e.g., Franzen, Haut, Rankin, & Keefover, 1995; Tombaugh & Hubley, 1997). The desired trend is for the proportion of correct responses to decrease (i.e. for the items to become more difficult) as the test progresses. For the 60-item BNT, we compared the difficulty index for each item to similar data from previously published research to help identify items that may be particularly problematic in the South African context.

Results

Sample characteristics

Participants ranged in age from 18–24 years ($M = 19.98 \pm 1.68$). They had completed between 12 and 17 years of education ($M = 13.31 \pm 1.12$). The modal annual income bracket for participants' suburb of residence was R153 601.00–R307 200.00.

60-Item Boston Naming Test: Performance, psychometric properties and item analyses

At T1, the sample's mean score was 51.51 (median = 52; mode = 55; SD = 5.33; and range = 35–59). This performance was significantly worse than that of normative samples of young adults from North America but was not significantly different from that of a comparable sample of highly educated young adults from New Zealand (Table 2). These results must be interpreted with caution, because the current BNT_{T1} scores were significantly non-normally distributed, Shapiro-Wilk test (59) = 0.92, $p = 0.001$, skewness = -0.93, kurtosis = 0.48.

Test-retest reliability was acceptable: T1 and T2 performance were significantly positively associated, Spearman's $\rho = 0.41$,

TABLE 2: Comparison of the current sample's Boston Naming Test 60-item performance to that of North American and New Zealand normative samples.

Study	N	Age		Education	Score	t	p	Cohen's d
		Range	M (SD)					
Farmer (1990) [†]	25	20–29	24.08 (2.53)	14.88 (1.67)	56.04 (3.60)	-6.54	< 0.001***	0.85
Tombaugh and Hubley (1997) [‡]	22	25–34	NR	NR	56.00 (2.90)	-6.48	< 0.001***	0.84
Barker-Collo (2001) [§]	58	17.5–25.25	20.2	14	52.5 (2.20)	-1.43	0.160	0.19
Mitrushina et al. (2005) [¶]	NR	25–29	NR	NR	55.71 (3.08)	-6.06	< 0.001***	0.79

Note: For the variable *Education*, data are *M* (SD) for the number of completed years.

NR, not reported in the original publication.

[†], Normative data for young adults from the United States. We chose to compare our sample's performance to that of participants in this study because it is one of the few studies that provides BNT norms for young adults.

[‡], Normative data for young adults from Canada. This study's BNT norms are widely used in clinical practice.

[§], Normative data for young adults (second-year and third-year university students) from New Zealand. We chose to compare our sample's performance to that of participants in this study because it is one of the few studies that provides BNT norms for highly educated young adults from outside of North America.

[¶], Meta-analytic norms. Fourteen North American studies (*N* = 1684), published between 1990 and 2003, were included in the meta-analysis.

***, *p* < 0.001; Bonferroni-corrected $\alpha = 0.05/4 = 0.0125$.

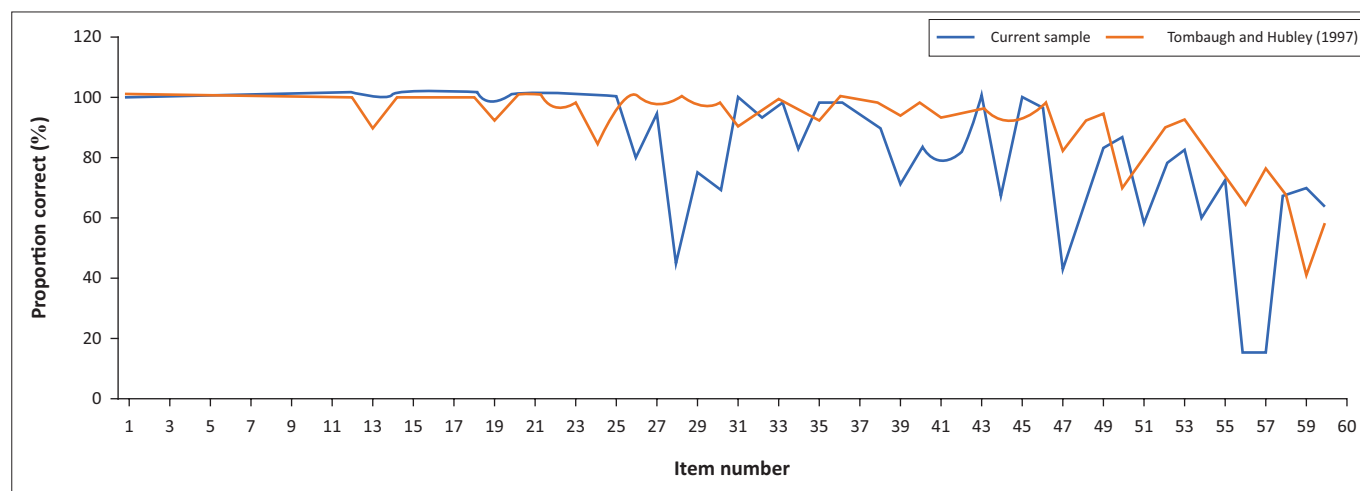


FIGURE 1: Item difficulty index for the current administration, at the first test occasion, of the standard 60-item Boston Naming Test. Data are proportion of correct responses made spontaneously or with stimulus cue for a sample of young English-speaking South African adults (*N* = 59). Comparative data (*N* = 219 English-speaking Canadian adults, age range 25–88, education range = 9–21 years) are from Tombaugh and Hubley (1997).

p = 0.001. Internal consistency reliability was better, however, with Cronbach's $\alpha_{T1} = 0.85$. Component variables with zero variances (viz., items 1–12, 14–18, 20–25, 31, 43 and 45) were not included in this analysis.

Figure 1 presents an item difficulty index based on the performance of the current sample at T1. Most of the easiest items (i.e. those to which 100% of participants responded correctly) are clustered at the beginning of the test. Although there is a roughly linear trend towards more difficult items at the end of the test, it is notable that the line is jagged, with more difficult items (e.g. 28 and 47) interspersed among much easier ones. A comparison of this item's difficulty index with that presented by Tombaugh and Hubley (1997) for their sample suggests that fully one-third of the 60 items might be regarded as culturally biased against South Africans (Table 3).

Boston Naming Test-South African Short Form: Performance, psychometric properties and item analyses

At T1, the sample's mean score was 13.97 (median = mode = 14; SD = 1.08; range = 11–15). This score was at least as good as the score they would have achieved on three other well-established 15-item short forms; in two of the

three cases, it was significantly higher (Table 4). These results must be interpreted with caution, however, because BNT-SASF_{T1} scores were significantly non-normally distributed, Shapiro-Wilk test (59) = 0.82, *p* < 0.001, skewness = -1.03, kurtosis = 0.49.

Analyses detected a significant positive association between BNT_{T1} and BNT-SASF_{T1} scores, Spearman's $\rho = 0.66$, *p* < 0.001. The estimate of test-retest reliability was confounded, however, because performance at T2 was better than that at T1 by at least one point in 77% of participants. Hence, performance at T1 was significantly negatively associated with that at T2, Spearman's $\rho = -0.39$, *p* = 0.037. Internal consistency reliability was poor, Cronbach's $\alpha_{T1} = 0.35$. Again, component variables with zero variances (viz., items 2, 7, 10, 15, 20, 22, 25, 31) were not included in this analysis.

Figure 2 presents an item difficulty index based on the performance of the current sample at T1. The trend for increasing errors as the test progresses is evident. Whereas all participants responded correctly to the first 8 items, there were increasing numbers of errors from items 11 through 15 (with the exception of item 12 [funnel], which appeared to be much more familiar to this sample than the items adjacent to it).

TABLE 3: Boston Naming Test item difficulty index: Current sample versus a North American sample.

Item	Current sample	Tombaugh and Hubley (1997) [†]
	(<i>N</i> = 59)	(<i>N</i> = 219)
13. octopus	98.3	90.0
24. seahorse	100.0	84.9
26. canoe	81.4	100.0
28. wreath	45.8	99.5
29. beaver	76.3	97.5
30. harmonica	69.5	96.8
31. rhinoceros	100.0	90.4
34. stilts	83.1	95.0
38. harp	89.9	97.3
39. hammock	71.2	94.1
40. knocker	83.1	97.7
41. pelican	78.0	92.7
42. stethoscope	81.4	95.0
44. muzzle	66.1	92.7
45. unicorn	100.0	91.3
47. accordion	42.4	81.7
48. noose	64.41	91.3
49. asparagus	83.1	93.6
50. compass	86.4	69.0
51. latch	57.6	80.8
52. tripod	79.7	89.5
53. scroll	83.1	92.7
54. tongs	59.3	84.5
56. yoke	15.3	63.0
57. trellis	17.0	77.2
59. protractor	69.5	39.7
60. abacus	64.4	57.5

Note: Data are shown only for those items where there was a difficulty difference of $\geq 5\%$. Items marked in boldface font are those on which performance of the current sample was $\geq 5\%$ worse than that of the North American sample.

[†], Sample of Canadian adults (age range = 25–88 years, education range = 9–21 years). This study's BNT normative data are widely used in clinical practice.

TABLE 4: Comparison of the current sample's Boston Naming Test-South African Short Form performance with that on other 15-item short forms (*N* = 59).

Test	Score	<i>t</i>	<i>p</i>	Cohen's <i>d</i>
CERAD [†]	13.27	-4.02	< 0.001**	0.52
Mack-SF4 [‡]	13.73	-1.55	0.128	0.20
Empirically derived short form [§]	12.32	-7.57	< 0.001***	0.99

BNT-SASF, Boston Naming Test-South African Short Form; CERAD Consortium to Establish a Registry for Alzheimer's Disease; Mack-SF4 = Mack Short Form 4.

[†], Morris et al. (1989); items 4, 7, 10, 13, 20, 23, 26, 29, 36, 39, 42, 45, 52, 55 and 58.

[‡], Mack et al. (1992); items 1, 2, 4, 5, 8, 10, 17, 18, 23, 26, 30, 35, 39, 36 and 54.

[§], Lansing et al. (1999); items 14, 21, 22, 24, 27, 32, 35, 37, 38, 39, 47, 49, 53, 57 and 58.

*, $p < 0.05$. **, $p < 0.01$; ***, $p < 0.001$; Bonferroni-corrected $\alpha = 0.05/3 = 0.017$.

Discussion

The Boston Naming Test has, for decades, been one of the most widely used neuropsychological tests (Rabin, Barr, & Burton, 2005; Rabin, Paolillo, & Barr, 2016). Despite its global reach and popularity, many of the test's items are heavily culture-bound. Hence, there is a high risk for misdiagnosis of naming deficits when the BNT is used to assess individuals outside of North America (Cruice et al., 2000; Tallberg, 2005).

The current study describes the development of, and preliminary psychometric properties for, a South African-adapted version of the BNT. Because local clinical conditions demand shorter and simpler forms of test administration, the BNT-SASF contains 15 items. These items were judged by a panel of practising neuropsychologists and community

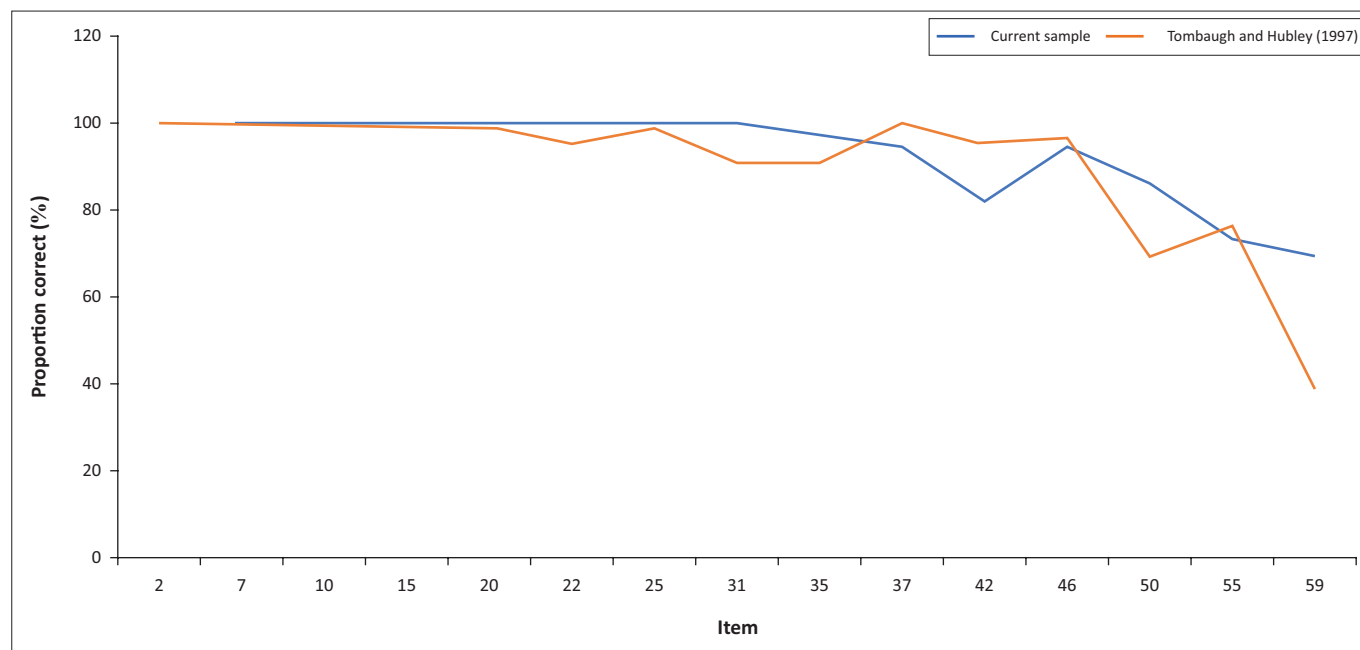
members to be culturally appropriate for local use. We administered the standard 60-item BNT, which incorporates the BNT-SASF, to a homogenous (English-fluent, high-SES, highly educated) sample of young adults. We reasoned that such a design, featuring the segment of the South African population that most closely matches North American normative samples, would allow us to avoid potentially confounding sociodemographic influences and to thus draw inferences about the basic utility of the BNT-SASF in this country.

Our analyses of BNT-SASF data suggested the instrument tests the same construct as other versions of the instrument. Most participants scored 14/15 at the first administration, a high level of performance that is consistent with North American samples administered different 15-item short forms (Fastenau et al., 1998; Lansing et al., 1999; Mack et al., 1992; Tombaugh & Hubley, 1997). Moreover, the performance of our participants on the 15 items comprising the BNT-SASF was better than their performance on the 15 items comprising other well-known short forms that were developed outside of South Africa and, therefore, without consideration of local cultural and contextual factors.

Boston Naming Test-South African Short Form scores were significantly positively associated with 60-item BNT scores, with the value of the correlation coefficient ($\rho = 0.66$) within the range reported in the literature on other 15-item short forms. That range spans values from 0.62 for the CERAD short form (Tombaugh & Hubley, 1997), through 0.74 for the Mack SF-4 (Fastenau et al., 1998), and up to > 0.95 for all Mack short forms (Franzen et al., 1995). The current correlation would have been stronger had performance on the 60-item BNT been as good as that on the short form. As discussed below, many of the 60 items proved to be relatively problematic for our participants and so their scores were relatively poor on the full instrument. Any discrepancy in favour of the BNT-SASF over the BNT might be interpreted as an indication of success in removing culturally biased items from the instrument.

Further evidence for the content validity of the BNT-SASF emerges from the item difficulty index created using the performance of the current sample. That index suggested that earlier items were relatively easy whereas later items were relatively difficult (with the last two items being the most difficult). This difficulty trend is what the BNT developers intended and the fact that performance on our 15-item version displays that trend is encouraging.

Although the internal consistency reliability of the BNT-SASF was quite low (Cronbach's $\alpha = 0.35$), the value of this estimate is in the same range as what Tombaugh and Hubley (1997) report for the CERAD short form and the Mack SF-4 ($\alpha = 0.36$ and 0.49 , respectively). It is unsurprising that these values are relatively low, given that the internal consistency of a test is strongly related to its length (i.e. tests with more items are typically more internally consistent;



Note: Item numbers correspond to the position of these 15 items within the standard 60-item Boston Naming Test. Data are proportion of correct responses made spontaneously or with stimulus cue for a sample of young English-speaking South African adults ($N = 59$). Comparative data ($N = 219$ English-speaking Canadian adults, age range 25–88, education range = 9–21 years) are from Tombaugh and Hubley (1997).

FIGURE 2: Item difficulty index for the current administration of the 15-item Boston Naming Test-South African Short Form.

Cohen & Swerdlik, 2018). This is one reason why some in this field prefer 30-item short forms over 15-item short forms (Williams et al., 1989).

A more prominent concern, however, is the relatively poor test–retest reliability ($\rho = -0.39$) of the BNT-SASF. As we note above, this value is influenced by the fact that most participants performed better at T1 than at T2 (perhaps as a result of carryover effects, specifically the administration of phonemic and multiple-choice cues at T1). Such poor test–retest reliability is not a typical feature of 15-item BNT short forms. For instance, Teng et al. (1989) reported a value of 0.90 over a 1-week interval for a sample of patients with AD. It is unclear, however, whether they followed standard administration procedures at both test occasions, as we did.

Our analyses of the current sample’s 60-item BNT data confirmed that the instrument’s inherent cultural biases make it unsuitable, in its original and unmodified form, for administration in South African clinical and research settings. We found, for instance, that the overall performance of our sample of English-fluent, high-SES, highly educated participants was significantly worse than that of comparable samples of young adults from North America and that the root of this performance difference was the difficulty our participants experienced on culturally bound items such as *wreath*, *beaver* and *yoke*. This result replicates those of numerous previous studies reporting on cross-cultural administration of the BNT (see, e.g., Barker-Collo, 2001; Worrall et al., 1995).

Regarding reliability of the 60-item BNT in the current sample, findings were mixed. Whereas internal consistency reliability ($\alpha = 0.85$) was within the range most commonly

cited as an acceptable value for this statistic (Cohen & Swerdlik, 2018), and was comparable to the coefficient ($\alpha = 0.78$) reported by Tombaugh and Hubley (1997), test–retest reliability ($\rho = 0.41$), although statistically significant, was relatively low compared to previous studies. For instance, Flanagan and Jackson (1997) reported a value of 0.90 over a 1–2-week interval for a sample of healthy older adults. Other studies of neurologically intact older adults suggest that this excellent test–retest reliability is maintained over much longer intervals (Mitrushina et al., 2005). Unfortunately, previous BNT investigations of healthy young adult samples do not provide reliability data. One possible reason for the relatively poor test–retest reliability in this sample is that our participants were farther away from ceiling effects at T1 than those in other samples, and improved significantly at T2 (again, perhaps as a result of carryover effects). Statistical comparison of T1 and T2 performance bears out this account, $t = -1.47$, $p = 0.15$, Cohen’s $d = 0.27$.

Limitations and directions for future research

The inferences we might draw from this study are limited by the size and nature of the sample. Compared with other studies that collected original data in developing BNT short forms (e.g. Fastenau et al., 1998; Graves et al., 2004), our sample size was smaller. Moreover, the sample was not representative of the national population, or even of the population of South African undergraduates (note that 45 of the 104 individuals we recruited did not meet our very strict eligibility criteria). However, the purpose of this study was not to collect nationally representative normative data, or to make generalised statements about the utility of the BNT-SASF. Instead, we intentionally recruited a homogeneous group of participants so as to avoid the confounding effects

of sociodemographic variables (e.g. age, education and home language) on performance, and then set out to show (as a first step in a meticulous process of psychometric investigation) that this new instrument is reliable and valid in a South African sample that is, broadly speaking, comparable to those used in most North American normative studies.

A second limitation is that, for at least two reasons, we cannot make definitive statements about the construct validity of the BNT-SASF. First, the magnitude of the correlation between BNT and BNT-SASF scores might be spuriously high as a result of method variance. Second, we did not administer independent tests of confrontation naming ability (e.g. the Naming Test of the Neuropsychological Assessment Battery; Yochim, Kane, & Mueller, 2009). We chose not to do so because all existing tests of that cognitive construct are of the same form (i.e. the participant views an image and is asked to identify the pictured object). Hence, comparative analyses of performance on the BNT and any of those tests runs the risk of being confounded by common method variance.

A third limitation is that, rather than collecting original cross-national data, we used historical data when comparing performance of the current sample to that of adults from other countries. Such historical comparisons are vulnerable to cohort effects and it is possible that we observed a minor instance of such effects here. For example, whereas 100% of our participants identified *unicorn* correctly, only 90% of Tombaugh and Hubley's (1997) sample and 83% of Barker-Collo's (2001) sample did so. The relative easiness of this item in the 2019 group might be attributed to the relatively more prominent place unicorns have in contemporary popular culture (Segran, 2017). One remedy for such circumstances is to engage in what Fernández and Abe (2018, p. 1) term 'simultaneous test development across multiple cultures'.

Follow-up studies of the BNT-SASF are already underway. In future articles, we will describe the psychometric properties of Afrikaans and isiXhosa versions of the instrument, report on how performance is influenced by age, education and SES, and investigate diagnostic validity in samples of healthy older adults and dementia patients. We encourage independent research groups to develop versions of the instrument appropriate for their own linguistic contexts, and to collaborate in collecting nationally representative and appropriately stratified normative data.

Summary and conclusion

Neuropsychological tests developed, standardised and normed in high-income countries of the global north often deliver misleading results when used outside of their sociocultural and linguistic context of origin (Howieson, 2019; Nell, 2000). This is especially true when the tests are used without critical consideration of cultural bias and cultural fairness, when construct validity in the local context has not been verified, or when locally appropriate normative data are not used. The need for cognitive tests that are reliable, valid, and culturally fair for use in South African

clinical and research settings is growing. Increasing numbers of neuropsychology trainees are entering the field. Increasing amounts of overseas grant money are being invested into South African-based neuroscience research but funded projects must use psychometrically sound instruments that are well known to international audiences.

Here, we described the development and psychometric assessment of a South African-adapted short form of the BNT. A key aspect of the BNT-SASF's value is that its items are drawn from the pool of items comprising the original test. This makes it a time- and cost-effective option on many levels (e.g. we did not have to curate an entirely new set of items, and those who already own the standard BNT will be able to use this modified short form without purchasing any new materials). These are particularly important considerations when one is operating in a resource-limited setting such as South Africa. Another advantage of this short form is that, unlike many other short forms that are developed via odd-even or split-half methods, this one was developed on an item-by-item basis, which lends itself to evaluation by item response theory (Pedraza et al., 2009). Our data suggest that the BNT-SASF demonstrates basic psychometric properties that are the equivalent of short forms developed elsewhere in the world. Moreover, it appears to measure the same construct as the full 60-item BNT while being less culturally biased.

Acknowledgements

Competing interests

The authors have declared that no competing interests exist.

Authors' contributions

K.G.F.T. designed the research, supervised the data analysis and wrote the first draft of the manuscript. L.B. contributed to research design, collected and analysed the data, wrote parts of the manuscript and approved the final version. C.Y.P. collected and analysed the data, wrote parts of the manuscript, contributed to manuscript preparation and approved the final version. H.L.F. led the research project and approved the final version to be published.

Funding information

This research was supported by the University of Cape Town (University Research Scholarship, awarded to L.B.) and the Stellenbosch University Strengthening Research Initiative Programme (Junior Research Fellowship, awarded to H.L.F.).

Data availability statement

Data are available upon request.

Disclaimer

The views and opinions expressed in this article are those of the authors and do not necessarily reflect the official policy or policy or position of any affiliated agency of the authors.

References

- Allegri, R.F., Villavicencio, A.F., Taragano, F.E., Rymberg, S., Mangone, C.A., & Baumann, D. (1997). Spanish Boston Naming Test norms. *The Clinical Neuropsychologist*, 11(4), 416–420. <https://doi.org/10.1080/13854049708400471>
- Balthazar, M.L.F., Cendes, F., & Damasceno, B.P. (2008). Semantic error patterns on the Boston Naming Test in normal aging, amnesic mild cognitive impairment, and mild Alzheimer's disease: Is there semantic disruption? *Neuropsychology*, 22(6), 703–709. <https://doi.org/10.1037/a0012919>
- Barker-Collo, S.L. (2001). The 60-item Boston naming test: Cultural bias and possible adaptations for New Zealand. *Aphasiology*, 15(1), 85–92. <https://doi.org/10.1080/02687040042000124>
- Barker-Collo, S.L. (2007). Boston Naming Test performance of older New Zealand adults. *Aphasiology*, 21(12), 1171–1180. <https://doi.org/10.1080/02687030600821600>
- Cohen, R.J., & Swerdlik, M.E. (2018). *Psychological testing and assessment: An introduction to tests and measurement* (9th edn.). New York: McGraw-Hill.
- Crowe, M., Clay, O.J., Martin, R.C., Howard, V.J., Wadley, V.G., Sawyer, P., & Allman, R.M. (2012). Indicators of childhood quality of education in relation to cognitive function in older adulthood. *The Journals of Gerontology: Series A*, 68(2), 198–204. <https://doi.org/10.1093/geron/gls122>
- Cruice, M.N., Worrall, L.E., & Hickson, L.M.H. (2000). Boston Naming Test results for healthy older Australians: A longitudinal and cross-sectional study. *Aphasiology*, 14(2), 143–155. <https://doi.org/10.1080/026870300401522>
- Farmer, A. (1990). Performance of normal males on the Boston Naming Test and the Word Test. *Aphasiology*, 4(3), 293–296. <https://doi.org/10.1080/02687039008249081>
- Fastenau, P.S., Denburg, N.L., & Mauer, B.A. (1998). Parallel short forms for the Boston Naming Test: Psychometric properties and norms for older adults. *Journal of Clinical and Experimental Neuropsychology*, 20(6), 828–834.
- Fernández, A.L., & Abe, J. (2018). Bias in cross-cultural neuropsychological testing: problems and possible solutions. *Culture and Brain*, 6(1), 1–35. <https://doi.org/10.1007/s40167-017-0050-2>
- Fernández, A.L., & Fulbright, R.L. (2015). Construct and concurrent validity of the Spanish adaptation of the Boston Naming Test. *Applied Neuropsychology: Adult*, 22(5), 355–362. <https://doi.org/10.1080/23279095.2014.939178>
- Fillenbaum, G.G., Huber, M., & Taussig, I.M. (1997). Performance of elderly white and African American community residents on the abbreviated CERAD Boston Naming Test. *Journal of Clinical and Experimental Neuropsychology*, 19(2), 204–210. <https://doi.org/10.1080/01688639708403851>
- Flanagan, J.L., & Jackson, S.T. (1997). Test-retest reliability of three aphasia tests: Performance of non-brain-damaged older adults. *Journal of Communication Disorders*, 30(1), 33–42. [https://doi.org/10.1016/S0021-9924\(96\)00039-1](https://doi.org/10.1016/S0021-9924(96)00039-1)
- Franzen, M.D., Haut, M.W., Rankin, E., & Keefover, R. (1995). Empirical comparison of alternate forms of the Boston Naming Test. *The Clinical Neuropsychologist*, 9(3), 225–229.
- Golden, Z., Bouvier, M., Selden, J., Mattis, K., Todd, M., & Golden, C. (2005). Differential performance of Alzheimer's and vascular dementia patients on a brief battery of neuropsychological tests. *International Journal of Neuroscience*, 115(11), 1569–1577. <https://doi.org/10.1080/00207450590957953>
- Graves, R.E., Bezeau, S.C., Fogarty, J., & Blair, R. (2004). Boston Naming Test short forms: A comparison of previous forms with new item response theory based forms. *Journal of Clinical and Experimental Neuropsychology*, 26(7), 891–902. <https://doi.org/10.1080/13803390490510716>
- Grima, R., & Franklin, S. (2016). A Maltese adaptation of the Boston Naming Test: A shortened version. *Clinical Linguistics & Phonetics*, 30(11), 871–887. <https://doi.org/10.1080/02699206.2016.1181106>
- Hawkins, K.A., & Bender, S. (2002). Norms and the relationship of Boston Naming Test performance to vocabulary and education: A review. *Aphasiology*, 16(12), 1143–1153. <https://doi.org/10.1080/02687030244000031>
- Howieson, D. (2019). Current limitations of neuropsychological tests and assessment procedures. *The Clinical Neuropsychologist*, 33(2), 200–208. <https://doi.org/10.1080/13854046.2018.1552762>
- 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>
- Kang, Y., Kim, H., & Na, D.L. (2000). Parallel short forms for the Korean–Boston naming test (K-BNT). *Journal of the Korean Neurological Association*, 18(2), 144–150.
- Kaplan, E.F., Goodglass, H., & Weintraub, S. (1978). *Boston Naming Test: Experimental edition*. Boston, MA: Boston University.
- Kaplan, E.F., Goodglass, H., & Weintraub, S. (1983). *The Boston Naming Test*. Philadelphia, PA: Lea & Febiger.
- Kaplan, E.F., Goodglass, H., & Weintraub, S. (2001). *The Boston Naming Test* (2nd edn.). Philadelphia, PA: Lippincott Williams & Wilkins.
- Katzef, C., Henry, M., Gouse, H., Robbins, R.N., & Thomas, K.G.F. (2019). A culturally fair test of processing speed: Construct validity, preliminary normative data, and effects of HIV infection on performance in South African adults. *Neuropsychology*, 33(5), 685–700. <https://doi.org/10.1037/neu0000539>
- 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>
- Kim, H., & Na, D.L. (1999). Normative data on the Korean version of the Boston Naming Test. *Journal of Clinical and Experimental Neuropsychology*, 21(1), 127–133. <https://doi.org/10.1076/jcen.21.1.127.942>
- Kiran, S., Cherney, L.R., Kagan, A., Haley, K.L., Antonucci, S.M., Schwartz, M., ... Simmons-Mackie, N. (2018). Aphasia assessments: A survey of clinical and research settings. *Aphasiology*, 32(Suppl 1), 47–49. <https://doi.org/10.1080/02687038.2018.1487923>
- Kohnert, K.J., Hernandez, A.E., & Bates, E. (1998). Bilingual performance on the Boston Naming Test: Preliminary norms in Spanish and English. *Brain and Language*, 65(3), 422–440. <https://doi.org/10.1006/brln.1998.2001>
- Lansing, A.E., Ivnik, R.J., Cullum, C.M., & Randolph, C. (1999). An empirically derived short form of the Boston Naming Test. *Archives of Clinical Neuropsychology*, 14(6), 481–487. [https://doi.org/10.1016/S0887-6177\(98\)00022-5](https://doi.org/10.1016/S0887-6177(98)00022-5)
- Leite, K.S.B., Miotto, E.C., Nitrini, R., & Yassuda, M.S. (2017). Boston Naming Test (BNT) original, Brazilian adapted version and short forms: Normative data for illiterate and low-educated older adults. *International Psychogeriatrics*, 29(5), 825–833. <https://doi.org/10.1017/S1041610216001952>
- Lezak, M.D., Howieson, D., & Loring, D. (2004). *Neuropsychological assessment* (4th edn.). New York: Oxford University Press.
- Lichtenberg, P.A., Ross, T., & Christensen, B. (1994). Preliminary normative data on the Boston Naming Test for an older urban population. *Clinical Neuropsychologist*, 8(1), 109–111. <https://doi.org/10.1080/13854049408401548>
- Lucas, J.A., Ivnik, R.J., Smith, G.E., Ferman, T.J., Willis, F.B., Petersen, R.C., & Graff-Radford, N.R. (2005). Mayo's older African Americans normative studies: Norms for Boston Naming Test, Controlled Oral Word Association, category fluency, animal naming, token test, wrat-3 reading, trail making test, stroop test, and judgment of line orientation. *The Clinical Neuropsychologist*, 19(2), 243–269. <https://doi.org/10.1080/13854040590945337>
- Lyu, J., & Burr, J.A. (2016). Socioeconomic status across the life course and cognitive function among older adults: An examination of the latency, pathways, and accumulation hypotheses. *Journal of Aging and Health*, 28(1), 40–67. <https://doi.org/10.1177/0898264315585504>
- Mack, W.J., Freed, D.M., Williams, B.W., & Henderson, V.W. (1992). Boston Naming Test: Shortened versions for use in Alzheimer's disease. *Journal of Gerontology*, 47(3), 154–158. <https://doi.org/10.1093/geronj/47.3.P154>
- Mitrushina, M., Boone, K.B., Razani, J., & D'Elia, L.F. (2005). *Handbook of normative data for neuropsychological assessment* (2nd edn.). New York: Oxford University Press.
- Morris, J.C., Heyman, A., Mohs, R.C., Hughes, J.P., Van Belle, G., Fillenbaum, G., ... the CERAD Investigators. (1989). The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. *Neurology*, 39(9), 1159–1165. <https://doi.org/10.1212/wnl.39.9.1159>
- Mosdell, J., Balchin, R., & Ameen, O. (2010). Adaptation of aphasia tests for neurocognitive screening in South Africa. *South African Journal of Psychology*, 40(3), 250–261. <https://doi.org/10.1177/008124631004000304>
- Neils, J., Baris, J.M., Carter, C., Dell'aira, A.L., Nordloh, S.J., Weiler, E., & Weisiger, B. (1995). Effects of age, education, and living environment on Boston Naming Test performance. *Journal of Speech, Language, and Hearing Research*, 38(5), 1143–1149. <https://doi.org/10.1044/jshr.3805.1143>
- Nell, V. (Ed.). (2000). *Cross-cultural neuropsychological assessment: Theory and practice*. Mahwah, NJ: Erlbaum.
- Patricacou, A., Psallida, E., Pring, T., & Dipper, L. (2007). The Boston Naming Test in Greek: Normative data and the effects of age and education on naming. *Aphasiology*, 21(12), 1157–1170. <https://doi.org/10.1080/02687030600670643>
- Pedraza, O., Graff-Radford, N.R., Smith, G.E., Ivnik, R.J., Willis, F.B., Petersen, R.C., & Lucas, J.A. (2009). Differential item functioning of the Boston Naming Test in cognitively normal African American and Caucasian older adults. *Journal of the International Neuropsychological Society*, 15(5), 758–768. <https://doi.org/10.1017/S1355617709990361>
- Rabin, L.A., Barr, W.B., & Burton, L.A. (2005). Assessment practices of clinical neuropsychologists in the United States and Canada: A survey of INS, NAN, and APA Division 40 members. *Archives of Clinical Neuropsychology*, 20(1), 33–65. <https://doi.org/10.1016/j.acn.2004.02.005>
- Rabin, L.A., Paolillo, E., & Barr, W.B. (2016). Stability in test-usage practices of clinical neuropsychologists in the United States and Canada over a 10-year period: A follow-up survey of INS and NAN members. *Archives of Clinical Neuropsychology*, 31(3), 206–230. <https://doi.org/10.1093/arclin/acw007>
- Robbins, R.N., Joska, J.A., Thomas, K.G.F., Stein, D.J., Linda, T., Mellins, C.A., & Remien, R.H. (2013). Exploring the utility of the Montreal Cognitive Assessment to detect HIV-associated neurocognitive disorder: The challenge and need for culturally valid screening tests in South Africa. *The Clinical Neuropsychologist*, 27(3), 437–454. <https://doi.org/10.1080/13854046.2012.759627>
- Roberts, P.M., Garcia, L.J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology*, 16(4–6), 635–645. <https://doi.org/10.1080/02687030244000220>
- Roebuck-Spencer, T.M., Glen, T., Puente, A.E., Denney, R.L., Ruff, R.M., Hostetter, G., & Bianchini, K.J. (2017). Cognitive screening tests versus comprehensive neuropsychological test batteries: A National Academy of Neuropsychology education paper. *Archives of Clinical Neuropsychology*, 32(4), 491–498. <https://doi.org/10.1093/arclin/acx021>
- 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>
- Segran, E. (2017). *The unicorn craze, explained*. Retrieved from <https://www.fastcompany.com/40421599/inside-the-unicorn-economy>.
- Statistics South Africa. (2011). *Census 2011*. Pretoria: Statistics South Africa.

- Strain, J.F., Didehban, N., Spence, J., Conover, H., Bartz, E.K., Mansinghani, S., ... Womack, K.B. (2017). White matter changes and confrontation naming in retired aging National Football League athletes. *Journal of Neurotrauma*, *34*(2), 372–379. <https://doi.org/10.1089/neu.2016.4446>
- Strauss, E., Sherman, E.M.S., & Spreen, O. (2006). *A compendium of neuropsychological tests: Administration, norms, and commentary* (3rd edn.). New York: 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>
- Teng, E.L., Wimer, C., Roberts, E., Damasio, A.R., Eslinger, P.J., Folstein, M.F., ... Henderson, V.W. (1989). Alzheimer's dementia: Performance on parallel forms of the Dementia Assessment Battery. *Journal of Clinical and Experimental Neuropsychology*, *11*(6), 899–912. <https://doi.org/10.1080/01688638908400943>
- Tombaugh, T.N., & Hubble, A.M. (1997). The 60-item Boston Naming Test: Norms for cognitively intact adults aged 25 to 88 years. *Journal of Clinical and Experimental Neuropsychology*, *19*(6), 922–932. <https://doi.org/10.1080/01688639708403773>
- Williams, B.W., Mack, W.J., & 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](https://doi.org/10.1016/0028-3932(89)90186-3)
- World Medical Association. (2013). World Medical Association Declaration of Helsinki; ethical principles for medical research involving human subjects. *The Journal of the American Medical Association*, *310*(20), 2191–2194. <https://doi.org/10.1001/jama.2013.281053>
- Worrall, L.E., Yiu, E.M.L., Hickson, L.M.H., & Barnett, H.M. (1995). Normative data for the Boston Naming Test for Australian elderly. *Aphasiology*, *9*(6), 541–551. <https://doi.org/10.1080/02687039508248713>
- Yochim, B.P., Kane, K.D., & Mueller, A.E. (2009). Naming Test of the Neuropsychological Assessment Battery: Convergent and discriminant validity. *Archives of Clinical Neuropsychology*, *24*(6), 575–583. <https://doi.org/10.1093/arclin/acp053>
- Zec, R.F., Burkett, N.R., Markwell, S.J., & Larsen, D.L. (2007). A cross-sectional study of the effects of age, education, and gender on the Boston Naming Test. *The Clinical Neuropsychologist*, *21*(4), 587–616. <https://doi.org/10.1080/13854040701220028>