Performance of French-speaking Quebec adults on the Boston Naming Test

Résultats du Boston Naming Test chez des adultes francophones de Québec

Patricia M. Roberts
Nathalie Doucet

Abstract

The Boston Naming Test (BNT) is an American test, widely used in Canada, in both English and French. This study examined the 60 stimuli of the BNT for name agreement in Quebec French as a first step towards developing scoring rules. Sensitivity of the test to age and education differences was also tested. Forty-five Quebec francophones (ages 50 to 79) with a mean of 8.6 years of education and little knowledge of English were asked to name all 60 pictures in French. The responses were scored after consulting 3 different dictionaries. There are 13 pictures with synonyms in French, and several additional items where it is unclear what names should be accepted. Nearly half the items are French-English cognates. Despite these problems with the stimuli, a 2 by 2 ANOVA (age by education) found significant effects (p <.008) for age and for education, indicating that the test, even with no cultural adaptation, may be sufficiently sensitive to detect different levels of naming ability. Future studies will focus on replacing some of the problematic pictures to achieve better name agreement and reduce the number of cognates.

Abrégé

Le Boston Naming Test est un test américain largement utilisé au Canada anglais et français. Cette étude a vérifié les 60 stimuli du test en français québécois, servant ainsi de première étape dans l'élaboration de règles de cotation. La sensibilité du test aux variables âge et éducation a aussi été testée. Quarante-cinq francophones du Québec (âgés de 50 à 79 ans) ayant une moyenne de 8,6 années d'éducation et une faible connaissance de l'anglais ont dû nommer 60 images en français. Les réponses ont été cotées après la consultation de trois dictionnaires différents. Il y a 13 images qui ont des synonymes en français et plusieurs autres images où il est difficile de déterminer quels noms devraient être acceptés. Près de la moitié des images sont des mots jumeaux en français et en anglais. En dépit de ces problèmes avec les stimuli, une analyse de variance à deux variables (âge par éducation) a trouvé un effet significatif pour l'âge et l'éducation (p < .008), indiquant que le test, même sans adaptation culturelle, pourrait être suffisamment sensible pour identifier différents niveaux de capacité de dénomination. Les études à venir seront axées sur le remplacement de certaines images problématiques, afin d'atteindre une meilleure concordance des noms et de réduire le nombre de mots jumeaux.

Patricia M. Roberts,
Faculty of Health Sciences,
University of Ottawa,
Ottawa, ON, Canada

Nathalie Doucet, Faculty
of Health Sciences,
University of Ottawa,
Ottawa, ON, Canada

The most widely used test of confrontation naming ability may be the Boston Naming Test (Kaplan, Goodglass & Weintraub, 2001). Published in the United States, and not designed as an international instrument, it has nonetheless enjoyed widespread use in North America, England, and Australia (Katz et al., 2000). There are studies of the test, with or without changes to the stimuli, in at least nine different languages. Garcia and Desrochers (1997) found that the Boston Naming Test (BNT) is often used to assess French-speaking patients in Canada. This is surprising, given that there are no published norms or even French scoring guidelines for the test. As this paper will show, scoring the BNT in French is far from simple.

THREE CHARACTERISTICS OF AN IDEAL TEST OF NAMING

Visually unambiguous: Ideally, a confrontation naming test is made up of pictures that are visually unambiguous (Snodgrass & Vanderwart, 1980). Clinical experience shows that some of the BNT drawings are frequently misidentified. The standardized cues to be given when a visual misidentification is made (e.g., saying *snake* for *pretzel*) or when the wrong part of the picture is named (e.g., *man* instead of *stilts*) compensate for this, to some degree. Rossion and Pourtois (2004) recommend that future editions of the BNT use coloured drawings as this would reduce the number of misidentifications.

Name agreement: Ideally, all pictures should have only one possible correct name (Lecours & Lhermitte, 1979; Lezak, 1995; 2004). Cuetos, Ellis, and Alvarez (1999) and Marquez et al., (2007) propose using 85% name agreement as a criterion for stimuli. A high level of name agreement for test stimuli is important for at least six reasons. Having a single possible correct answer:

1. makes scoring easy and quick, which is important in clinical settings;

2. enhances inter-rater reliability. If the patient changes clinicians over the course of their recovery, each new clinician can administer the test, confident that their scoring is neither stricter nor looser than that of their colleagues. Similarly, across published studies, scores can be compared without idiosyncratic scoring rules acting as a confounding factor;

3. makes it possible to give phonemic cues. If there are two or three possible names for a given picture, it is impossible to know which one the person being tested is trying to find (Bender, Martin Garcia, & Barr, 2010; Roberts & Bois, 1998; 1999). Giving a cue that does not match the word being searched for, or the word on the tip of the tongue of the person being tested could lead him/ her to mistakenly conclude that the planned answer was wrong;

4. reduces response time and allows response times to be compared across items and experimental groups. Mitchell (1989) found that response time was slower for items with more than one possible name;

5. allows analysis based on the properties of that single answer. If there are two possible names for a picture (e.g., *harmonica* or *mouth organ*), it is impossible to analyse results based on factors such as familiarity, frequency and word length. All of these factors influence response time and accuracy of naming (e.g., Rochefort & Williams, 1962; Dent, Johnson, & Humphreys, 2008). Word frequency affects naming success in many aphasic adults (Nickels & Howard, 1995);

6. allows analysis based on likely age of acquisition. Hodgson and Ellis (1998) have found that age of acquisition influences naming speed and accuracy. These effects persist in aphasic speakers (e.g., Cuetos, Aguado, Izura, & Ellis, 2002; Nickels & Howard, 1995). For a recent review see Johnson and Barry's recent work (2006).

Sensitivity: A test should accurately discriminate between different levels of ability and allow for testing a range of ages or levels without encountering floor or ceiling effects. It has been shown that the BNT does this quite well in studies of English-speaking North Americans, when a range of ages and education levels and adequate power levels are used (Neils, et al., 1995; Randolph, Lansing, Ivnik, Cullem, & Hermann, 1999; Welch, Doineau, Johnson, & King, 1996; Zec, Burkett, Markwell, & Larsen, 2007a). Some studies with low power and /or fairly narrow ranges of age or education show little or no correlation between BNT scores and education or age (e.g., Nicholas, Brookshire, MacLennan, Schumacher, & Porrazzo, 1989). Neurologically intact adults younger than 70 years with at least a high school education tend to score near the maximum (54 to 57 out of 60), in studies in the USA and Canada (Randolph et al.,1999; Tombaugh & Hubley, 1997; Zec et al., 2007a, 2007b) but this is not the target group for this test.

In a growing body of studies in different languages, the BNT has accurately discriminated between groups with different levels of education (e.g., Allegri, et al.,1997; Patricacou, Psallida, Pring, & Dipper, 2007; Rami et al., 2008) and those with and without language impairments

(e.g., Cheung, Cheung, & Chan, 2004; Marquez de la Plata et al., 2008; Tsolaki, Tsantali, Lekka, Kiosseoglu, & Kazis, 2003).

TESTS ACROSS LANGUAGES AND CULTURES

When a naming test is used to assess speakers of languages other than the one it was written in, two additional characteristics to consider are cultural appropriateness and cognates.

Cultural appropriateness

Many authors have pointed out the need to consider cultural factors when tests "travel" from one country or language to another (e.g., Ardila, 2007; Ardila, Rosselli, & Puente, 1994; Garcia & Desrochers, 1997). If pretzels are not a common snack food in New Zealand or Greece, and if testing shows a large discrepancy between the percent correct in different cultural groups/countries, then this item should be either replaced, or retained but moved to a better place in the sequence from easy to difficult items. Some thought and testing are needed to ensure that either the replacement item is of roughly the same difficulty as the original item it is replacing and that multiple replacements balance each other to maintain the range of difficulty of the original test samples.

It is important not to overextend the notion of cultural appropriateness as applied to tests (see also Roberts, 2001, 2008). Lapointe-Goupil, Everett, Rousseau, Giguère, Laplante, and Keller (2004), in the only published study of

Table 1

Examples of studies with different versions of the BNT in different languages and countries

Items in bold font were found to be problematic in the present study

Language/country	Version	Reference
English – Canada	English – original	Tombaugh & Hubley, 1997
English – New Zealand	English – original but recommend replacing **pretzel** and beaver	Barker-Collo, 2007
French – Canada	Original but noted culturally inappropriate items	Lapointe-Goupil et al., 2004
French – Switzerland	Found low sensitivity to education for the items: abacus, beaver, cactus, dart, hanger, igloo, pelican, seahorse, **stethoscope, scroll,** unicorn, **latch**, protractor, tripod, volcano, whistle, **yoke.** Replaced these in final, shortened version of the test for use in French	Thuillard-Colombo & Assal, 1992
Swedish	Original	Tallberg, 2005
Greek – Greece	Replaced 4 items: **pretzel, door knocker, stethoscope, scroll** with, respectively, a kind of cake, mailbox, blood pressure instrument, ancient Greek column	Patricacou, 2007
Spanish – Argentina, Spain	Original Spanish version replaces 15/60 items: whistle; saw; toothbrush; mushroom; hanger; wheelchair; **pretzel**; seahorse; wreath; **escalator; door knocker**; asparagus; **latch; scroll**; trellis. However, the more recent Neuronorma project (Peña – Casanova et al., 2009) kept all original items except **pretzel** (Spain)	The original Spanish version is used in Argentina (Allegri et al., 1997) and Spain (Rami et al., 2007); a newer version was created by Peña – Casanova et al. (2009)
Cantonese – Hong Kong	Selected 30 items from original test	Cheung, Cheung, & Chan, 2004

the BNT with Quebec French-speaking adults, state that sphinx and sea horse should be replaced as these items are "not very representative of Quebec context" (p. 204: "peu représentatifs du contexte québécois"). By the same logic, *camel, rhinoceros,* and *pyramid* would also have to be replaced. These things are not part of American or English Canadian culture either, but it is reasonable to expect educated American and Canadian adults to know what they are. Similarly, Thuillard-Colombo and Assal (1992) and Lapointe-Goupil et al. (2004) both object to noose because it has negative associations. One assumes this would be true in all cultures. However, this is not a valid, culturally-based reason to delete it from the test.

Aside from the published Spanish version of the test (Garcia-Albea, Sanchez-Bernardos, & del Viso-Pabon, 1986), researchers and clinicians generally use the published English version, asking patients to respond in their own language. Some studies in European countries have recommended deleting or replacing some items as culturally inappropriate. Other studies (see Table 1) have used the 60 item test as is and found significant effects for age and education (e.g., Marien, Mampaey, Vervaet, Saerens, & De Deyn, 1998; Tallberg, 2005). The original 60 pictures were used, but regionalisms and synonyms were scored as correct for 15 pictures in Italian (Riva, Nichelli & Devoti, 2000) and 37 in Dutch (Marien et al., 1998).

Some authors have chosen to make quite massive changes. For example, Kim and Na (1999) replaced 49 of the 60 pictures with no published data to justify the replacements. In more measured approaches, the first step has been to field test the 60 pictures to explore name agreement and whether the item is familiar to people in that specific country. For Greek adults, Patricacou and colleagues (2007) asked 10 well-educated adults to name "culturally specific" items. Four items which 8 of the 10 volunteers failed to name correctly were replaced. The name agreement for these (newly drawn) pictures was then checked, informally, with 10 new adult participants. In selecting replacement items, the authors tried to stay within the same general category. This approach contrasts with that of Thuillard-Colombo and Assal (1992), assessing French-speaking Swiss adults. The authors created an 85 item experimental version of the test but then, prior to testing, removed nine items they felt were culturally inappropriate. These included the items *wheelchair* and *noose*, both seen as too negative based on the authors' judgement, without any testing.

The starting premise of the current study is that before deciding that a particular item is problematic, the original test should be given, and the results scrutinized for name agreement, difficulty level, and sensitivity. Intuition is not a sufficient basis for modifying a test.

Cognates

Ideally, a naming test used to assess bilingual speakers should have a low number of cognate words in the pairs of languages tested. Cognates are words with the same meaning and very similar phonology and/or spelling in two languages. A test with as few cognates as possible is desirable because:

1. cognate words are processed differently in bilingual speakers (Costa, Santeseban & Caño, 2005; de Groot & Nas, 1991);

2. cognates are easier to name than non-cognates (Gollan, Fennema-Notestine, Montoya & Jernigen, 2007; Lemhofer, Dijkstra, & Michel, 2004; Roberts & Deslauriers, 1999);

3. in clinical practice, it is sometimes difficult to tell if a patient is pronouncing a word in the target language or not, if the two words are very similar in the two languages;

4. if the words are cognates, it is possible to "make up" the equivalent word in the other language and end up with a correct answer.

The goals of the present, exploratory study were to assess the name agreement in French for the 60 BNT items and to determine if the unmodified BNT is sensitive to the effects of age and education in a sample of elderly speakers of Quebec French.

METHOD

Participants

Forty-five adults (28 women and 17 men) aged between 50 and 79 years old volunteered to participate in the study. They were from three towns in Quebec (Drummondville, Trois-Rivières and Lévis) where 70% to 75% of citizens report being only able to speak French (Statistics Canada, 2006). Most participants responded to notices and announcements made in church halls and community centres. Some were located via friends and acquaintances of the second author. Overall, education levels were low (mandatory education laws were not in effect during their youth). For the entire group, the mean number of years of education was 8.6 years (4 to 22 years of education for the younger group under 65 years and 3 to 14 years of education for the group aged 65 and over). Ten participants in the younger age group and 13 in the older group had less than 9 years of education. The mean education levels for each age group are listed in Table 2. The most common occupations were housewife (12), textile or clothing factory worker (8), construction worker or

Table 2

Age and education levels of participants

		Age				
		Below 65 years		**65 or more**		**All**
		Mean (SD)	**Range**	**Mean (SD)**	**Range**	**Mean Age**
Education	**< 9 years**	58.2 (3.8) *n* = 10	51 to 63	71.3 (5.01) *n* = 13	65 to 78	65.6 (7.98) *n* = 23
	≥ 9 years	56.5 (4.35) *n* = 13	51 to 64 *n* = 9	73.0 (3.84)	66 to 79	63.2 (9.26) *n* = 22

carpenter (4) and mechanic (3). Seven participants were office workers (2 teachers, 1 secretary, 1 receptionist, 2 pastoral care workers, 1 government office worker). The other occupations (one of each) were butcher, clerk/ housewife, seamstress/ housewife, day labourer, farmer, janitor, paper factory worker, landlord, mailman, jewelry craftsman, and steelworker.

In analyses, participants were divided into two groups for level of education (< 9 years versus 9 or more years of education), and two age groups (50—69 versus 70-79 years old) as shown in Table 2. The mean number of years of education was significantly different for the two education groups at 6.00 versus 11.2 [t(27.55, corrected for unequal variances) = 7.07, p < .001]. The two groups based on age also significantly differed in mean age: 57.2 versus 72.0 [t(43) = - 11.44, p < .001]. The younger group mean number of years of education was 9.7 (range 4 to 22, SD 4.1) while the older group mean years of education was 7.3 (range 3 to 14, SD 2.5) The difference in education levels between the younger and older groups was not significant [t (43) = 1.04, p = 0.306].

Language background was documented using a relatively short questionnaire administered in French, by a native speaker of Quebec French. All participants stated that French was their native language and the language used at least 95% of the time in their daily lives (work, home, TV/ computer, social life). The only other language knowledge reported was English. None of the participants reported learning any English before age 10. Of the 13/ 45 who stated that they "spoke another language", 11 were in the group with higher level of education. Of these, 11/ 13 claimed to speak, understand, read, and write French "much better" than English (the options were: same, a little better, much better). The other two

reported speaking French (1) or understanding French (1) "a little better than" English. None of the 13 reported using English more than 5% of the time at work or at home. Only 3 reported ever used English more than 5% of the time. One woman reported speaking English to a friend for one year when she was 11 or 12 years old. Another woman spent one year in the USA when she was in her early twenties, but reports never using English at home or at work. Thus, even those reporting some knowledge of English rarely used it.

All participants stated they had no communication problem, memory disorders, head injury or stroke and no neurologic or psychiatric troubles. None reported a history of alcohol or substance abuse.

Each volunteer was seen individually in his or her home or in a local community centre. All 60 pictures were presented in order. If no answer was given after 20 seconds or if an answer was a clear misidentification of the picture (e.g., *snake* for *pretzel*), the researcher gave a semantic cue. No phonemic cues were given. The total correct was answers with no cue plus those following the semantic cue.

Scoring

It was necessary to develop a list of acceptable answers to score the test. In English, all items have (ostensibly) only one correct, one-word name, if the person is told that all names are single words (to avoid composites like park bench, tennis racquet, Christmas wreath; however, see Roberts, Garcia, Desrochers, & Hernandez, 2002 and Zec, Markwell, Burkett, & Larsen, 2005, for regional synonyms for some items in English). This is not the case in French. For some items, the only possible answer is more than one word (e.g., brosse à dents). For others, there is more than one correct, widely used name: Seahorse

can be either *hippocampe* or *cheval de mer*. Snail can be colimaçon or escargot. These items were scored as correct if either name was given.

Some items have names widely used in some regions that are not correct, standard French. These regionalisms were accepted if at least one Canadian dictionary of French listed them as correct (Bélisle, 1979; Boulanger, 1992; Proteau, 1991). The expected names, synonyms in standard French, and regional variants are presented in Table 3.

If the participant gave more than one answer (it could be a _____ or a _____), they were asked to pick one and if they could not choose, the second answer was the one we scored. If a participant named the object and then said something like "But I don't know", we scored the name. The correct response plus an adjective was scored as correct, since participants were not told to give only one word answers, as noted above (e.g., *grosse maison* for *maison*). Errors in pronunciation were accepted if only one phoneme was incorrect and the resulting response was not another real word. For example, *captus* for *cactus* was scored as correct. All scoring was checked by both authors and any disagreements resolved through discussion and careful review of the 3 dictionaries of Canadian French chosen as the basis for scoring decisions.

Table 3

Translations and alternate responses for English and French
All names listed in this table were counted as correct responses

English (alternate or synonym)	French (synonym)	Non-standard, regional
1. bed	lit	
2. tree	arbre	
3. pencil	crayon	
4. house	maison	
5. whistle	sifflet	
6. scissors (pair of scissors)	ciseaux (paire de ciseaux)	
7. comb	peigne	
8. flower	fleur	
9. saw (hand saw)	scie	égoine
10. toothbrush	brosse à dents	
11. helicopter	hélicoptère	
12. broom	balai	
13. octopus	pieuvre	
14. mushroom (toadstool)	champignon	
15. hanger (clothes or coat hanger)	cintre (support)	
16. wheelchair	chaise roulante (fauteuil roulant)	
17. camel	chameau	
18. mask	masque	
19. pretzel	bretzel (pretzel)	
20. bench	banc	
21. raquet	raquette	
22. snail	escargot (calimaçon)	
23. volcano	volcan	
24. sea horse	hippocampe (cheval de mer)	
25. dart	fléchette (dard)	
26. canoe	canot (canoë)	
27. globe	**globe terrestre**	mappemonde, boule terrestre
28. wreath	couronne	
29. beaver	castor	
30. harmonica (mouth organ)	harmonica	musique à bouche, ruine babines

31. rhinoceros (rhino)	rhinocéros	
32. acorn	gland	
33. igloo	igloo	
34. stilts	échasses	
35. dominoes	dominoes	
36. cactus	catus	
37. escalator	**escalier roulant**	
38. harp	harpe	
39. hammock	hamac	
40. knocker (door knocker)	**marteau (heurtoir)**	
41. pelican	pélican	
42. stethoscope	stéthoscope	
43. pyramid	pyramide	
44. muzzle	muselière	
45. unicorn	licorne (unicorne)	
46. funnel	entonnoir	
47. accordion (squeeze box)	accordéon	
48. noose	**corde de potence (nœud coulant)**	
49. asparagus (asparagus spears)	asperge	
50. compass	compas	
51. latch	**loquet**	**barrure, clenche**
52. tripod	trépied	
53. scroll	**manuscrit**	**parchemin**
54. tongs	**pincettes (pinces)**	
55. sphinx	sphynx	
56. yoke	joug	
57. trellis	treillis	tuteur
58. palette	**palette**	planche à/ pour peintre planche à peinture; planche à toile
59. protractor	rapporteur d'angles	
60. abacus	boulier (boulier compteur)	

Bold indicates items that are very difficult to score in French. See text for alternate names.
Responses with adjective added (e.g., park bench, tennis racquet, Halloween mask) were accepted as correct and are not listed in this table.

RESULTS

Name Agreement

In the English BNT, disregarding items often given two-word names (tennis racquet, coat hanger, Halloween mask…), there are only 3 items with more than one correct name in Canadian English: mushroom/ toadstool (drawing is ambiguous), harmonica/ mouth organ, and accordion/ squeeze box. Thirteen items were found to have two possible names in standard French, as listed in most dictionaries. These are *saw, harmonica, hanger, wheelchair, pretzel, snail, seahorse, dart, canoe, globe, unicorn, tongs,* and *abacus.* Table 4 shows the percent correct for each item, including synonyms.

In addition, there were 6 items where scoring was very difficult, even after consulting dictionaries (*escalator, noose, latch, scroll, yoke,* and *palette*). This was partly due to a lack of consensus among native speakers on the names and partly due to the nature of French. Many object names are composite nouns, with *à* or *de* linking words that make up a single name: toothbrush – brosse à dents (brush for teeth) and cheval de mer - seahorse. This means that when people say "corde à _____" for *noose,* it is difficult to set a dividing line between "appropriate, multi-word name" and circumlocution or description given because the person does not know the name of this picture. Few participants gave either of the dictionary labels for this item (*noeud coulant* or *corde de potence*). *Palette* is also

difficult to score for the same reason. *Planche* has many meanings in French, so the name of this particular object cannot be a single word. It is unclear at what point the many variations become descriptive, circumlocution type errors: *planche à peinture, planche de peintre, planche pour peintre, planche pour (la/les) peinture(s)*.

Scoring is made more difficult by the tendency of Canadian French to accept the English word, in some instances, or the English word, given a French twist. Thus *escalator* pronounced as if it were a French word is seen by many people as wrong (based on an informal check with native French speakers enrolled in speech-language pathology) yet it was found in one of the three chosen dictionaries and, therefore, scored as correct. It is unsatisfactory that the choice of dictionary could lead to different scoring decisions on some responses. The need to distinguish between commonly used and lexically correct labels requires that a clinician make scoring judgements that are not required in English. Cumulatively, these could change an individual's score by 6 or more out of 60. These scoring dilemmas reduce the inter-rater reliability to unacceptable levels, in the absence of any published BNT scoring guidelines in French.

In Table 4, errors in one phoneme are included in the total correct for the relevant word. These errors were (number of participants who made the error): hélico_tère 1; bratzel 1; fretzel 1; ninocéros 3; rhonocéros 1; captus 9; tactus 1; _ téthoscope 2; mususlière 1; entonnoi_ 2. Of the 22 instances of these slips of the tongue, only 4 were made by participants over 70 years old. All 9 captus/cactus errors were by adults under age 70.

Table 4

Percent of participants correctly naming each picture (due to rounding, totals and subtotals do not always add up) % of answers for each possible correct name shown in parentheses

1. bed	100
2. tree	98
3. pencil	98
4. house	89
5. whistle	98
6. scissors	100
7. comb	100
8. flower	91
9. saw - (scie 56) (égoine 44)	100
10. toothbrush	98
11. helicopter	87
12. broom	91
13. octopus	87
14. mushroom	98
15. hanger (cintre 49) (support 51)	100
16. wheelchair (chaise roulante 71) (fauteil roulant 18)	89
17. camel	80
18. mask	87
19. pretzel (pretzel 49) (bretzel 7)	56
20. bench	100
21. raquet	93
22. snail (escargot 24) (calimaçon 36)	60
23. volcano	89
24. sea horse (hippocampe 16) (cheval de/s mer/s 31)	47
25. dart (fléchette 13) (dard 51)	64
26. canoe (canot 73) (canoë 7) (chaloupe 7)	87
27. globe (globe terrestre 76) (mappemonde 11) (boule terrestre 4)	91
28. wreath	98
29. beaver	84
30. harmonica (harmonica 16) (musique à bouche 78)	94
31. rhinoceros	75
32. acorn	78
33. igloo	71
34. stilts	40
35. dominoes	62
36. cactus	67
37. escalator (escalier roulant 58) (escalier mobile 7) (escalier électrique 4) (marches roulantes 4) (escalator 4)	78
38. harp	76
39. hammock	89
40. knocker (marteau 2) (heurtoir 4)	6

41. pelican	31	54. tongs (pincettes 4)	71
42. stethoscope	34		
43. pyramid	76	55. sphinx	29
44. muzzle	55	56. yoke (joug 22) (attelage 13) (licou 4) (collier 4) (harnais 2)	45
45. unicorn (licorne)	42		
46. funnel	85		
47. accordion	100		
48. noose (corde de potence 2) (noeud coulant 13) (câble avec noeud coulant 2) (corde de pendu 5) (câble à pendu 2) (câble de pendaison 4)	28	57. trellis (treillis 18) (tuteur 7)	29
49. asparagus	62	58. palette (palette 22) (planche à peintre 7) (planche pour peintre 2) (planche à peinture 2) (planche à toile 2)	34
50. compass	69		
51. latch (loquet 2) (barrure 58) (clenche 9) (verrou 4)	73		
		59. protractor	18
52. tripod	62	60. abacus (boulier 47) (boulier compteur 9) (abaque 2)	58
53. scroll (parchemin 36) (manuscrit 2)	38		

Table 5

Mean number of correct names

	Years of education					
	≤ 9 years		9 years or more		Total	
	Standard	Dialect	Standard	Dialect	Standard	Dialect
50-64 years old						
Mean	38.2	43.9	43.54	48.8	41.2	46.6
S.D.	(5.75)	(6.05)	(4.03)	(3.92)	(5.45)	(5.42)
Range	29 to 49	31 to 52	36 to 51	40 to 56	29 to 51	31 to 56
n	10	10	13	13	23	23
64-79 years old						
Mean	33.3	37.6	39.11	44.2	35.7	40.3
S.D.	(9.05)	(8.13)	(4.43)	(6.28)	(7.92)	(8.02)
Range	18 to 54	22 to 50	31 to 44	33 to 52	18 to 54	22 to 54
n	13	13	9	9	22	22
All groups (n=45)						
Mean	35.43	40.35	41.7	46.9	38.5	43.56
S.D.	(8.02)	(7.8)	(4.7)	(5.4)	(7.2)	(7.5)
Range	18 to 54	22 to 52	31 to 51	33 to 56	18 to 54	22 to 56

Note:
Standard: accepted standard French translation, with or without adjective; Dialect: also accepted regionalisms and colloquialisms
The four people with the highest education levels (22, 16, 15, and 14 years of education) obtained relatively high scores (49, 49, 48 and 42, respectively). These were the only participants with post-secondary education. However, the three highest scores on the test (56, 52 and 52) were obtained by individuals with 12, 10 and 9 years of education, respectively.

SENSITIVITY TO AGE AND EDUCATION

The number of correct answers varied with age (higher scores for the younger groups) and with education (higher scores for those with more than 9 years of schooling). As Table 5 shows, this was true for both methods of scoring. Accepting regionalisms and other variants raised the mean group scores by 4 to 7 points.

STATISTICAL ANALYSIS.

After checking for homogeneity of variance (Levene tests for each proposed comparison, all $p > .1$), two analyses of variance were done (2 x 2 : education x age), one using standard French and the other with regionalisms accepted (see Table 6). Regardless of the scoring method (accepting only standard French responses or also accepting regionalisms) there were significant effects for education and for age, both with large effect sizes (Cohen's d). There were no significant age by education interactions in either ANOVA.

DISCUSSION

Name Agreement on Stimuli

The BNT stimuli, as a set, failed to meet minimum levels of name agreement in French. The 85% standard discussed in the introduction was not met for 17 of the pictures. If the test were developed in French, no one would select so many items with more than one name. Multiple possible names make it difficult to score the test in a reliable way. Table 3 is not a definitive list; there are too many items where it remains unclear which answers should be considered correct. Depending on which dictionary of Canadian French is used to guide decisions about alternative names for pictures, the scoring rules would change.

Interpretation of test results may be particularly problematic for French-English bilingual speakers. Depending on the level of resemblance needed to be considered cognates, the BNT has between 23 and 30 French-English cognates. In some cases, one possible French name is a cognate and the other is not (e.g., *dart* could be *dard* or *fléchette*). The large number of cognate words helps those with some knowledge of English to score higher on the test in French by guessing. Roberts and Deslauriers (1999) showed that this strategy (dubbed the Bilingual Chipmunk Strategy) is used by both aphasic and non-aphasic, bilingual adults. The issue of cognates has been raised elsewhere (Roberts et al., 2002) and will be explored in future studies where participants have higher levels of proficiency and greater use of both languages. Given the limited English knowledge and patterns of use by the participants in the current study, the cognate status of possible picture names likely had little to no impact on scores.

Table 6

ANOVA (2 x 2) for Age and Education

	df	F	p	Partial eta²	Power
Accepting only Standard French responses					
Age	1	5.98	.019	.127	.66
Education	1	8.54	.006	.172	.81
Age by Education	1	0.015	.904	.000	.052
Error	41	(39.82)			
Accepting Regionalisms					
Age	1	8.127	.007	.165	.80
Education	1	9.123	.004	.182	.84
Age by Education	1	0.209	.650	.005	.073
Error	41	(39.56)			

Note: Values enclosed in parentheses represent Mean square errors.

Replacing 4 to 6 items on the test would simplify the scoring. This approach would be similar to that followed for the Greek and Australian adaptations (Cruise, Worrall, & Hickson, 2000; Patricacou, Psallida, Pring, & Dipper, 2007), while being somewhat less radical than some of the Spanish adaptations. Candidates for replacement, based on poor name agreement and scoring ambiguities are *tongs, latch, yoke, palette, noose,* and *escalator.* The fact that two of these are French-English cognates (palette and, for some, *escalator*) provides a further reason for replacing them. *Door knocker* was correctly named by 97% of English speaking Canadians, with at least 12 years of schooling (Tombaugh & Hubley, 1997) but by only 6% of the present sample. Any revised version of the test will need to consider extreme differences like this in balancing the overall difficulty level (and hence the sensitivity) of the test. A study of a larger and more representative sample of Canadian French speakers (on the 60 BNT items and on possible replacement items) will be necessary before making decisions about replacements. Some additional items might be replaced to reduce the number of French-English cognate items in the test.

Age and Education Effects

Despite these scoring and stimulus problems, the mean score on the BNT was significantly affected by both education and age, no matter how the test was scored. Thus, our results are consistent with those of other studies (see Table 7) and they indicate a degree of validity for the test, in French. Furthermore, the mean scores are similar to those of English-speaking Americans with similar, low levels of education. There were phonemic errors on 7 different words which we scored as correct. Scoring these as errors would slightly lower the overall group mean for the younger group but it would remain similar to that of adults with little education in other studies. In their study of American English speakers, Neils et al. (1995) said "according to our results for non-institutionalized subjects, the cut-off for the sixth to ninth grade education group should be 35; and the cut-off for the two higher educational groups should be 40" (p. 1148). By this standard, the unmodified test administered in French correctly identified 21/22 participants with 9 or more years education as scoring within normal limits. Of the 21 participants with 6 to 8 years of education, only

Table 7

Comparison with other studies (the sub-group most similar to the present study)

	Age range	n	Education (years)	Mean BNT Score/60	S.D.
French-Quebec					
Roberts & Doucet (standard)	50-79	45	8.6	38.5	7.2
(dialect)				43.6	7.5
Lapointe-Goupil 2004	n/a	17	8.8	44.4	6.6
	(mean 69.9)				
Mariën et al. 1998 (Dutch)	55-91	92	<10	49.6	6.4
Patricacou 2007 (Greek)	20 to >71	26	0 to 6	33.9	8.5
American-low education					
Neils et al. (1995)(USA-Eng)	65-74	12	6 to 9	47.6	6.1
	75-84	19	6 to 9	42.8	10.9
	65-74	xx	10 to 12	53.0	6.6
Nicholas et al.,1989 (USA-Eng)	40-78		13.9	54.5	
Welch & al., 1996	80+	20	<12	40.7	11.3
	75-79	17	<12	36.6	10.9

Note: Age and education are in years; n/a means not available; standard = correct French ; Dialect = regionalisms also accepted.

2 scored below this American cut-off score. Both had 6 years education and their scores were of 31 and 22 /60. Since there was no screening for cognitive ability, it is unclear whether these low scores (especially the 22/60) are due to the BNT stimuli or to neurocognitive problems.

The low level of education and the fact that only 3 of the 45 completed university make it hard to generalise results to the broader population of Francophones in Quebec. With the introduction of mandatory education laws in the 1960s (requiring that all children remain in school until age 16), groups composed of people born since the mid-1950s will have considerably more education than those in this study. Therefore, it will be important to replicate this study with a larger and more representative sample of adults, in Quebec and in other regions where French is spoken (especially regions in Ontario and New Brunswick) before drawing final conclusions about the validity of the BNT in French.

Education effects on BNT scores have been found in various languages in most studies with adequate power (e.g., Marien et al., 1998; Neils et al., 1995; Patricacou et al., 2007; Thuillard-Colombo & Assal, 1992; Welch et al., 1996). As in other studies (e.g., Randolph, Lansing, Ivnik, Cullum, & Hermann, 1999; Tombaugh & Hubley, 1997; Welch et al., 1996), education influenced scores slightly more than age. Given that only 3 of the 45 participants completed university, and most did not complete high school, the range of education levels was not ideal. With a broader and more balanced distribution, there would likely be a stronger education effect. The fact that a significant education effect was found, despite the small *n* and a skewed distribution in education levels, testifies to the potential sensitivity and validity of the BNT in French.

Nevertheless, much of the variance in scores remains unaccounted for. Education and age together only explain 34% of the variability. Future studies will be necessary to identify further sources, including level of bilingualism, reading habits, and the frequency of occurrence of each possible name in French compared to English.

Despite the problems with the stimuli, it is both encouraging and puzzling that the unchanged BNT was sensitive to differences in education and age. Significant effects were found with both lexically strict scoring rules and with broader rules that accepted regional and colloquial names. Lapointe-Goupil et al. (2004) dismissed the BNT as "culturally biased" because items such as *sphinx* are not part of Quebec culture. The authors proposed a new, 36 item test using pictures from the Snodgrass and Vanderwart set (Bellefleur, 2001). However, their own results show that the French BNT (with no details provided as to how it was scored, and with small groups) was sensitive to differences in performance

between a control group and a group of adults with clinical depression. These results should encourage rather than discourage the use of the BNT in Quebec.

There are a number of advantages to using the same test across different languages, with adaptations where necessary to ensure roughly equal difficulty and sensitivity. Although other naming tests have been developed in French, including the Masters thesis of Bellefleur (see Lapointe-Goupil et al., 2004) and a European test (Metz-Lutz, Kremin, Deloche, Hannequin, Ferrand, Perrier, et al., 1991), these tests are not easy to obtain (especially now that the Snodgrass and Vanderwart pictures are copyrighted and must be purchased). The age of acquisition for some names and the name itself may not always be the same for the different varieties of French spoken in Europe and in Canada. Using the BNT (modified or not) facilitates comparisons across languages, especially for studies of bilingualism and bilingual language acquisition and loss.

Before recommending or discarding the BNT for use with French-speaking Canadians, more study is needed.

ACKNOWLEDGMENTS

REFERENCES

Acevedo, A., & Loewenstein, D. A. (2007). Performance on the Boston Naming Test in English-Spanish bilingual older adults: Some considerations. *Journal of the International Neuropsychological Society, 13*(2), 212-214.

Allegri, R. F., Mangone, C.A., Villavicencio, A. F., Rymberg, S., Taragano, F. E., & Baumann, D. (1997). Spanish Boston Naming Test norms. *The Clinical Neuropsychologist, 11*, 416-420.

Ardila, A. (2007). Toward the development of a cross-linguistic naming test. *Archives of Clinical Neuropsychology, 22*, 297-307.

Ardila, A., Rosselli, M., & Puente, A. E. (1994). *Neuropsychological evaluation of the Spanish speaker.* New York: Plenum.

Barker-Collo, S. (2007). Boston Naming Test performance of older New Zealand adults. *Aphasiology, 21*, 1171-1180. doi:10.1080/02687030600821600

Bellefleur, S. (2001). Développement d'un outil de dénomination adapté à la population francophone nord-américaine : normes préliminaires. Unpublished manuscript. Laval University.

Bélisle, L.-A. (1979). *Dictionnaire Nord-Américain de la langue Française.* Montréal: Beauchemin.

Bender, H. A., Martin Garcia, A., Barr, W. B. (2010). An interdisciplinary approach to neuropsycho-logical test construction: Perspectives from translation studies. *Journal of the International Neuropsychological Society, 16*, 227 – 232 . doi:10.1017/S13556177709991378227

Boulanger, J.-C. (1992). *Dictionnaire québécois d'aujourd'hui : langue française, histoire et géographie.* St-Laurent, QC. Dicorobert.

Cheung, R.W., Cheung, M. C., & Chan, A. S. (2004). Confrontation naming in Chinese patients with left, right, or bilateral damage. *Journal of the International Neuropsychological Society, 10*, 46–53. doi: 10.10170S1355617704101069

Costa, A., Santesteban, M., & Caño, A. (2005). On the facilitatory effects of cognate words in bilingual speech production. *Brain and Language, 94*, 94-103. doi:10.1016/j.bandl.2004.12.002

Cruise, 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, 12*, 143-155.

Cuetos, F., Aguado, G., Izura, C., & Ellis, A. W. (2002). Aphasic naming in Spanish: Predictors and errors. *Brain and Language, 82,* 344-365.

Cuetos, F., Ellis, A., & Alvarez, B. (1999). Naming times for the Snodgrass and Vanderwart pictures in Spanish. *Behavior Research Methods, Instruments, & Computers, 31,* 650–658.

de Groot, A. M. B., & Nas, G.L. (1991). Lexical representation of cognates and noncognates in compound bilinguals. *Journal of Memory and Language, 30,* 90–123.

Dent, K., Johnston, R. A., & Humphreys, G. W. (2008). Age of acquisition and word frequency effects in picture naming: A dual-task investigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*(2), 282-301.

Garcia, L. J. & Desrochers, A. (1997). L'évaluation des troubles du langage et de la parole chez l'adulte francophone. *Journal of Speech Language Pathology and Audiology, 21,* 271-293.

Garcia-Albea, J. E., Sanchez-Bernardos, M. L., & del Viso-Pabon, S. (1986). Test de Boston para el diagnóstico de la afasia: Adaptación española [Boston Naming Test for aphasia diagnosis: Spanish version]. *de la afasia y de trastornos relacionados* [Assessment of aphasia and related disorders] (2nd ed., pp. 129–198). Madrid, Spain: Editorial Medica Panamericana.

Gollan, T. J., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T.J. (2007). The bilingual effect on Boston Naming Test performance. *Journal of the International Neuropsychological Society, 13,* 197-208.

Hodgson, C. & Ellis, A. W. (1998). Last in, first to go: age of acquisition and naming in the elderly. *Brain and Language, 64,* 146-163.

Johnson, R.A. & Barry, C. (2006). Age of acquisition and visual processing. *Visual Cognition, 13,* 789-845. doi:10.1016/S0093-934X(03)00262-1

Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test.* Philadelphia: Lea & Febiger.

Katz, R., Hallowell, B., Code, C., Armonstrong, E., Roberts, P., Pound, C., & Katz, L. (2000). Multinational comparison of aphasia management practices. *International Journal of Language and Communication Disorders, 33,* 303-314.

Kim, H. & Na, D.L. (1999). Normative data on the Korean version of the Boston Naming Test. J*ournal of Clinical and Experimental Neuropsychology, 21,* 127-133.

Lapointe-Goupil, R., Everett, J., Rousseau, F., Giguère, V., Laplante, L., & Keller, E. (2004) Étude comparative entre le Boston Naming Test et le Test de Dénomination Bellefleur. *Revue québécoise de psychologie, 25*(2), 203-212.

Lecours, A. R. & Lhermitte, F. (1979). *L'aphasie.* Montréal : Les Presses de l'Université de Montréal.

Lemhöfer, K., Dijkstra, T., & Michel, M-C. (2004). Three languages, one ECHO: Cognate effects in trilingual word recognition. *Language and Cognitive Processes, 19,* 585-611.

Lezak, M., D., Howieson, D.B., & Loring, D.W.(2004). Neuropsychological assessment, 4th edition.New York: Oxford University Press.

Lezak, M. (1995). *Neuropsychological assessment.* 3rd edition. New York: Oxford University Press.

Mariën, P., Mampaey, E., Vervaet, A., Saerens, J., & De Deyn, P. P. (1998). Normative data for the Boston Naming test in native Dutch-speaking Belgian elderly. *Brain and Language, 65,* 447-467. doi:10.1006/brln.1998.2000

Marquez de la Plata, M., Vicioso, B., Hynan, L., Evans, H.M., Diaz-Arrastia, R., Lacritz, L., & Cullum, C.M. (2007). Development of the Texas Spanish Naming Test: a test for Spanish speakers. *The Clinical Neuropsychologist, 22,* 288-304. doi: 10.1080/13854040701250470

Metz-Lutz, M. N., Kremin, H., Deloche, G., Hannequin, D., Ferrand, I., Perrier, D., Quint, S., Dordain, M., Bunel, G., Cardebat, D., Larroque, C., Lota, A. M., Pichard, B. & Blavier, A. (1991). Standardisation d'un test de dénomination orale: contrôle des effets de l'âge, du sexe et du niveau de scolarité chez les sujets adultes normaux. *Revue de Neuropsychologie, 1*(1), 73-95.

Mitchell, D. B. (1989). How many memory systems? Evidence from aging. *Journal of Experimental Psychology: Learning, Memory and Cognition, 15,* 31-49.

Neils, J., Baris, J. M., Dell'aira, A. L., Nordloh, S. J., Weiler, E., & Weisger, B. (1995). Effects of age, education and living environment on Boston Naming test performance. *Journal of Speech and Hearing Research, 38,* 1143-1149.

Nicholas, L. E., Brookshire, R. H., MacLennan, D. L., Schumacher, J. G., & Porrazzo, S. A. (1989). Revised administration and scoring procedures for the Boston Naming Test and norms for non-brain-damaged adults. *Aphasiology, 3,* 569-580. doi: 10.1080/02687038908249023

Nickels, L., & Howard, D. (1995). *Aphasic naming- what matters? Neuropsychologia, 33,* 1281-1303. doi:10.1016/0028-3932(95)00102-9

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,* 1157-1170. doi: 10.1080/02687030600670643

Peña-Casanova et al. (2009). Spanish multicenter normative studies (NEURONORMA project): Norms for Boston Naming Test and Token Test. *Archives of Clinical Neuropsychology, 24,* 343-354. doi:10.1093/arclin/acp039

Ponton, M.O., & Ardila, A. (1999). The future of neuropsychology with Hispanic populations in the United States. *Archives of Clinical Neuropsychology, 14,* 565-580.

Proteau, L. (1991). *Le français populaire au Québec et au Canada.* Boucherville, Qc : Publications Proteau.

Rami, L., Serradell, M.,Bosch, B., Caprile, C., Sekler, A., Villar, A., Canal, R., & Molinuevo, J.L. (2008). Normative data for the Boston Naming Test and the Pyramids and Palm Trees Test in the elderly Spanish population. *Journal of Clinical and Experimental Neuropsychology, 30,* 1-6. doi: 10.1080/13803390701743954

Randolph, C., Lansing, A.E., Ivnik, R.J., Cullum, C.M., & Hermann, B.P. (1999). Determinants of confrontation naming performance. *Archives of Clinical Neuropsychology, 14,* 489-496. doi:10.1016/S0887-6177(98)00023-7

Riva, D., Nichelli, F., & Devoti, M. (2000). Developmental aspects of verbal fluency and confrontation naming in children. *Brain and Language, 71,* 267-284. doi:10.1006/brln.1999.2166

Roberts, P.M. (2001). Aphasia assessment and treatment in bilingual and multicultural populations. In R. Chapey (Ed.). *Language intervention strategies in adult aphasia.* 4th edition (pp. 208-232). Baltimore: Lippincott, Williams & Wilkins.

Roberts, P. M. (2008). Issues in assessment and treatment for bilingual and culturally diverse patients. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 245-275). Baltimore: Lippincott Williams & Wilkins.

Roberts, P. M. & Bois, M. (1998, June). Picture-name agreement for French-English bilingual adults. Poster presented at TENNET XII (Theoretical and Experimental Neuropsychology – Neuropsychologie expérimentale et théorique). Montreal.

Roberts, P. M., & Bois, M. (1999). Picture name agreement for French-English bilingual adults. *Brain and Cognition, 40,* 238-241.

Roberts, P.M. & Deslauriers, L. (1999). Picture naming of cognate and non-cognate nouns in bilingual aphasia. *Journal of Communication Disorders, 32,* 1-14. doi:10.1016/S0021-9924(98)00026-4

Roberts, P. M., Garcia, L. J., Desrochers, A., & Hernandez, D. (2002). English performance of proficient bilingual adults on the Boston Naming Test. *Aphasiology, 16,* 635-645. doi:10.1080/02687030244000220

Rochefort, G. & Williams, M. (1962). Studies in the development and breakdown of the use of names. *Journal of Neurology, Neurosurgery and Psychiatry, 25,* 222-233.

Rossion, B., & Pourtois, G. (2004). Revisiting Snodgrass and Vanderwart's object pictorial set: the role of surface detail in basic level object recognition. *Perception, 33,* 217-236. doi:10.1068/p5117

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity and visual complexity. *Journal of Experimental Psychology: Learning, Memory and Cognition, 6,* 174-215.

Statistics Canada 2006 Census: http://www12.statcan.gc.ca/census-recensement/2006/dp-pd/prof/92-591/index.cfm?Lang=E

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

Thomas, J. C., Fozard, J. L. & Waugh, N. C. (1977). Age-related differences in naming latency. *American Journal of Psychology, 90,* 499-509.

Thuillard-Colombo, F. & Assal, G. (1992). Adaptation française du test de dénomination de Boston versions abrégées. *Revue Européenne de Psychology Appliquée, 42*(1), 67-71.

Tombaugh, T. N. & Hubley, A. M.(1997). The 60 item Boston Naming Test: norms for cognitively intact adults aged 25 to 88 years. J*ournal of Clinical and Experimental Neuropsychology,19,* 922-932. doi: 10.1080/01688639708403773

Tsolaki, M., Tsantali, E., Lekka, S., Kiosseoglou, S., & Kazis, S. (2003). Can the Boston Naming Test be used as clinical tool for differential diagnosis in dementia. *Brain and Language, 87,* 185-186. doi:10.1016/S0093-934X(03)00262-1

Welch, L. W., Doineau, D., Johnson, S. & King, D. (1996). Educational and gender normative data for the Boston Naming Test in a group of older adults. *Brain and Language, 53,* 260-266.

Zec, R. F., Burkett, N. R., Markwell, S. J., Larsen, D. L. (2007a). A cross-sectional study of the effects of age, education, and gender on the Boston Naming Test. *The Clinical Neuropsychologist. 21,* 587-616. doi:

10.1080/13854040701220028

Zec, R. F., Burkett, N.R., Markwell, S.J., & Larsen, D.L. (2007b). Normative data stratified for age, education, and gender on the Boston Naming Test. *The Clinical Neuropsychologist. 21,* 617-637. doi: 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,* 716-726. doi: 10.10170S1355617705050897

AUTHOR'S NOTE

Correspondence (in French or English) should be sent to Patricia M. Roberts, Audiology and Speech-Language Pathology, 451 Smyth Road, Room 3071, University of Ottawa, Ottawa, ON., K1H 8M5, Canada. E-mail: proberts@uottawa.ca ▶