



## Research Article

# The copenhagen cross-linguistic naming test (C-CLNT): Development and validation in a multicultural memory clinic population

T. Rune Nielsen<sup>1</sup> , Bernadette Unmack Grollenberg<sup>1</sup>, Signe Pertou Ringkøbing<sup>1</sup>, Maria Özden<sup>2</sup>, Brendan Weekes<sup>3,4</sup> and Gunhild Waldemar<sup>1,5</sup>

<sup>1</sup>Department of Neurology, Danish Dementia Research Centre, Copenhagen University Hospital–Rigshospitalet, Copenhagen, Denmark, <sup>2</sup>Department of Neurorehabilitation and Spinal Cord Injuries, The Neuroscience Centre, Copenhagen University Hospital–Rigshospitalet, Hvidovre, Denmark, <sup>3</sup>Faculty of Education, The University of Hong Kong, Hong Kong, <sup>4</sup>The University of Melbourne, Parkville, Australia and <sup>5</sup>Department of Clinical Medicine, Faculty of Health and Medical Sciences, University of Copenhagen, Copenhagen, Denmark

### Abstract

**Objective:** Despite recent advances in cross-cultural neuropsychological test development, suitable tests for cross-linguistic assessment of language functions are not widely available. The aims of this study were to develop and validate a brief naming test, the Copenhagen Cross-Linguistic Naming Test (C-CLNT), for the assessment of culturally, linguistically, and educationally diverse older adult populations in Europe. **Method:** The C-CLNT was based on a set of standardized color drawings. Items for the C-CLNT were selected by considering name agreement and frequency across five European and two non-European languages. Ambiguities in some of the selected items and scoring criteria were resolved after pilot testing in 10 memory clinic patients. The final 30-item C-CLNT was validated by verifying its psychometric properties in 24 controls and 162 diverse memory clinic patients with affective disorder, mild cognitive impairment, and with dementia. **Results:** The C-CLNT had acceptable scale reliability (coefficient alpha = .67) and good construct validity, with moderate to strong correlations with traditional language tests ( $r = .42 - .75$ ). Diagnostic accuracy for dementia was good and significantly better than that of the Boston Naming Test (areas under the curve of .80 vs .64,  $p < .001$ ), but was poor for mild cognitive impairment. Only 3% of the variance in C-CLNT test scores was explained by immigrant background, while 6% was explained by age and years of education. In comparison, these proportions were 34 and 22% for the BNT. **Conclusions:** The C-CLNT has promising clinical utility for cross-linguistic assessment of naming impairment in culturally, linguistically, and educationally diverse older adults.

**Keywords:** language impairment; naming; dementia; mild cognitive impairment; immigrant; cross-cultural comparison; cultural diversity

(Received 23 March 2023; final revision 27 June 2023; accepted 28 June 2023)

### Introduction

With increasing demographic aging, migration, and globalization, there is a pressing need for standardized neuropsychological tests suited for diverse older populations (Nielsen, 2022). A certain degree of diversity has always been present in European countries, but cultural, language, and educational diversity has increased greatly over the last seven decades due to increasing mobility within the European Union as well as immigration from countries outside Europe (Van Mol & De Valk, 2016). Although immigration patterns differ between countries, the largest immigrant groups across Europe originate from other European countries, the Middle East, North Africa, and South Asia, followed by groups of Sub-Saharan African and Latin American origin (Nielsen et al., 2022). Despite recent advances in cross-cultural neuropsychological test development, suitable tests for cross-linguistic assessment of language functions are not widely available (Franzen et al., 2020). Naming impairment is frequent across several neurocognitive disorders, including stroke (RELEASE Collaborators,

2021), traumatic brain injury (Strain et al., 2017), and a variety of neurodegenerative disorders (Grossman et al., 2004). Most dementia syndromes are associated with naming impairment due to varying degrees of semantic memory impairment, impaired lexical retrieval, or impaired visual perception, depending on the subtype (Taler & Phillips, 2008). For instance, anomia is one of the core features of semantic dementia (Gorno-Tempini et al., 2011) and although memory impairment is generally the core feature of early AD, anomia is another common feature, especially as the disease progresses (McKhann et al., 2011). Thus, assessment of naming impairment is standard in most neuropsychological assessments and is typically measured with confrontation naming tests (Strauss et al., 2006) – for example, the Boston Naming Test (BNT; Kaplan et al., 2001).

Performance on confrontation naming tests is influenced by several linguistic and cultural variables. The difficulty level of the individual items depends on factors such as word frequency, familiarity, age of acquisition, length, visual complexity, and name

**Corresponding author:** T. R. Nielsen; Email: [rune.nielsen@rh.regionh.dk](mailto:rune.nielsen@rh.regionh.dk)

**Cite this article:** Nielsen T.R., Grollenberg B.U., Ringkøbing S.P., Özden M., Weekes B., & Waldemar G. (2023) The copenhagen cross-linguistic naming test (C-CLNT): Development and validation in a multicultural memory clinic population. *Journal of the International Neuropsychological Society*, 29: 911–921, <https://doi.org/10.1017/S1355617723000437>

© INS. Published by Cambridge University Press 2023. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

and image agreement (Ivanova & Hallowell, 2013), and these vary between cultures and languages (Ardila, 2007; Bertola & Malloy-Diniz, 2018; George & Mathuranath, 2007). For instance, items such as a *pretzel*, *beaver*, and *asparagus* included in the BNT may be familiar to people living in North America and parts of Europe but less familiar or virtually unknown in other cultural contexts (Franzen et al., 2023). In contrast, *abacus* is a difficult item to name in North America but relatively easier in China as abacuses are more common there (Gollan et al., 2012).

Although the BNT has been adapted to several languages and has had widespread clinical and research applications (Maruta et al., 2011; Rabin et al., 2016), cross-cultural research has shown that the items included in the BNT are suboptimal for assessing confrontation naming abilities in culturally, linguistically, and educationally diverse populations. More specifically, several studies have shown large differences in BNT performance between ethnoracial groups in the United States (US) (Baird et al., 2007; Boone et al., 2007), and lower performances in bilinguals (Gollan et al., 2007; Kohnert et al., 1998; Roberts et al., 2002) and second-language speakers (Stålhammar et al., 2022), even after controlling for differences in level of education and other demographic variables. Furthermore, BNT performance has been shown to be influenced by level of education (Strauss et al., 2006). This may reflect increasing vocabulary and exposure to a wider range of concepts with increasing levels of education. In addition, it has also been suggested that people with limited education and literacy may find it difficult to process items presented as black-and-white line drawings but perform better when the same items are presented in colored photographs or drawings (Reis et al., 2006; Reis et al., 2001). While some of these challenges may be overcome by using language and culture-adapted versions of the BNT and applying relevant normative adjustments, this clearly does not solve all issues.

Use of language and culture-specific confrontation naming tests and norms derived for native speakers have limited feasibility in most European memory clinics, in which patients may differ widely in their cultural, linguistic, and educational characteristics (Franzen et al., 2023; Nielsen, 2022). Therefore, to develop a reliable and valid solution for clinical practice it is important to develop confrontation naming tests with potential applicability across diverse cultural and language groups (Franzen et al., 2022).

During the last two decades, there have been several efforts to develop cross-linguistic naming tests. However, many of these efforts have resulted in tests with an inadequate balance between cross-linguistic properties and sensitivity to naming impairment. Thus, tests such as Body Part Naming from the Cross-Cultural Neuropsychological Test Battery (Dick et al., 2002), the Cross-Linguistic Naming Tests (Ardila, 2007), and Picture Naming from the European Cross-Cultural Neuropsychological Test Battery (Nielsen et al., 2018) have good cross-linguistic properties but poor sensitivity to milder language impairment in patients with Alzheimer's disease (AD) and other dementia disorders due to ceiling effects (Abou-Mrad et al., 2017; Araujo et al., 2020; Ardila, 2007; Dick et al., 2002; Gálvez-Lara et al., 2015; Nielsen et al., 2019b). In contrast, the abbreviated version of the Multilingual Naming Tests (MINT; Ivanova et al., 2013) has high sensitivity to milder language impairment in AD but is biased toward more highly educated white English speakers in the US (Franzen et al., 2023; Li et al., 2022; Paplikar et al., 2022; Stasenکو et al., 2019). More recent efforts include the Indian Council of Medical Research Picture Naming Tests (ICMR-PNT; Paplikar et al., 2022) and the Naming Assessment in Multicultural Europe

(NAME; Franzen et al., 2023). Both instruments have shown promising clinical utility for cross-linguistic assessment. However, the ICMR-PNT may be less useful outside the Indian subcontinent as some items, such as *tabala* (a musical instrument), are culture-specific, and the 60-item NAME is rather long, taking up to 20 minutes to administer, which impedes its clinical utility in a busy clinical setting.

Building on these efforts, our aims were to develop and validate a brief cross-linguistic naming test for assessment of culturally, linguistically, and educationally diverse older adult populations in Europe. We compared the diagnostic accuracy of the novel confrontation naming test and traditional language tests in a diverse memory clinic population as well as the psychometric properties of the tests. Items included objects as well as pictured actions since action naming impairment is also diagnostic of dementia (Parris & Weekes, 2001). The rationale for the study design, comparing patients with dementia, mild cognitive impairment (MCI), affective disorder, and subjective cognitive decline (SCD), is that these conditions are common differential diagnoses in patients referred to neuropsychological evaluation in memory clinic settings.

## Method

### Participants

Patients were recruited from the Copenhagen University Hospital Memory Clinic at Rigshospitalet, which is a multidisciplinary outpatient clinic based in the Department of Neurology. For this study, patients with immigrant background referred for neuropsychological evaluation as part of their diagnostic assessment were selectively included between June 2021 and June 2022. Patients with a majority ethnic Danish background were consecutively recruited in the same period. As described below, all patients in the clinic are assessed with cognitive screening tests as part of the basic diagnostic assessment. One of the criteria for referral to more comprehensive neuropsychological evaluation (approximately 2 hours) in the clinic is a Mini-Mental State Examination (MMSE; Folstein et al., 1975) or Rowland Universal Dementia Assessment Scale (RUDAS; Storey et al., 2004) score  $\geq$  22 at the initial visit in the clinic, but patients with lower MMSE or RUDAS scores may also be referred if necessary (e.g., in the case of patients with aphasia). In total, 169 patients who completed both the Copenhagen Cross-Linguistic Naming test (C-CLNT) and BNT were included in the study. Exclusion criteria included severe psychiatric symptoms and a diagnosis other than dementia, MCI, affective disorder, or SCD. In total, seven patients were excluded (two diagnosed with sequelae from traumatic brain injury, two with sequelae from stroke, two with epilepsy and psychiatric disorder, and one with atypical Parkinsonian disorder), resulting in a final sample of 162 patients.

All patients had an extensive diagnostic assessment including an interview with the patient and (when possible) an informant; a neurological, physical, and psychiatric examination including cognitive assessment with the MMSE and Addenbrooke's Cognitive Examination (Mathuranath et al., 2000) or the RUDAS and Multicultural Examination (Nielsen et al., 2019a) in case of cultural, linguistic and/or educational barriers; laboratory screening with blood tests and electrocardiography; and structural brain imaging with magnetic resonance imaging and/or computerized tomography. Further investigations, including functional imaging with [ $^{18}$ F]FDG-PET, amyloid imaging with [ $^{11}$ C]PIB-PET, and/or dopamine transporter imaging with [ $^{18}$ F]FE-PE2I PET, cerebrospinal fluid biomarker analysis, and

comprehensive psychiatric or neuropsychological evaluation were performed on clinical indication. Diagnoses were based on evidence from all clinical and investigational results, except the C-CLNT, applying the 5<sup>th</sup> edition of the Diagnostic and Statistical Manual of Mental Disorders (American Psychological Association, 2013) criteria for dementia, and diagnostic research criteria for specific dementia subtypes (Gorno-Tempini et al., 2011; McKeith et al., 2017; McKhann et al., 2011; Rascovsky et al., 2011; Sachdev et al., 2014), MCI (Winblad et al., 2004), and SCD (Jessen et al., 2014). Affective disorder (e.g., depression, anxiety, post-traumatic stress disorder) was diagnosed by applying the 10<sup>th</sup> edition of the International Classification of Diseases criteria (World Health Organization, 1993). Professional interpreters provided by interpretation services were freely available to patients during diagnostic assessments, including neuropsychological evaluation, when considered necessary.

Also, 24 cognitively intact participants aged 60 years or older were recruited from local general practice clinics and through the social networks of multicultural and multilingual researchers. Participants were assessed in their private homes, in the Copenhagen University Hospital Memory Clinic, or in another suitable location, depending on their preference. Participation was voluntary and without any economic incentive. All cognitively intact participants were living independently, reported no significant memory problems, psychiatric or neurological disorders, or substance abuse, and scored  $\geq 24/30$  points on the MMSE or  $\geq 23/30$  points on the RUDAS, and  $\leq 6/15$  points on the 5/15-item Geriatric Depression Scale (Weeks et al., 2003).

### Procedure

All participants underwent an approximately two-hours clinical assessment, in which medical and demographic data were collected and neuropsychological tests, including the C-CLNT, were administered. All assessments were made by specialists in neuropsychology. The comprehensive neuropsychological evaluation in the Copenhagen University Hospital Memory Clinic is based on a flexible assessment approach, meaning that a standardized, fixed set of neuropsychological tests covering the main cognitive domains is given to most patients with some flexibility to add or subtract tests given the specific referral question (Nielsen et al., 2022). The applied tests generally come from the international literature, but locally developed tests are also used. In case of cultural, linguistic, and/or educational barriers, patients are mainly assessed with tests from the European Cross-Cultural Neuropsychological Test Battery (Nielsen et al., 2019b; Nielsen et al., 2018). Participants with immigrant background were assessed in their primary language, either by a multilingual neuropsychologist (in Danish, English, Kurdish, or Turkish;  $n = 30$ ) or through interpreter-mediated assessment ( $n = 30$ ).

Demographic data collected at the clinical assessment included data on age, sex, years of education, country of origin, and mother tongue. For participants with an immigrant background, years of residence in Denmark were calculated by subtracting the year of the assessment from the year of immigration. Also, mother tongue was classified as a European or non-European language and cultural distance between the original culture and Danish culture was calculated using the Kogut and Singh Index (KSI; Kogut & Singh, 1988).

All participants were asked about any vision or hearing impairment and were assessed using their hearing aids or prescribed glasses when this was confirmed.

Results for the C-CLNT were compared with three traditional language tests: BNT (Full and 15-item version) and Category Fluency. In all tests, correct responses in any language were accepted.

The BNT contains 60 black-and-white line drawings ranked according to difficulty. In this study, the Danish adaptation of the BNT was used, in which the original items were ranked according to difficulty in a sample of older Danish typical participants (Jørgensen et al., 2017), using a discontinuation rule of six consecutive failures. The score is the number of correct responses, including responses after semantic cues. The score range is 0–60 points.

Scores for the abbreviated 15-item version of the BNT (BNT-15) introduced by the Consortium To Establish a Registry for Alzheimer's Disease (Morris et al., 1988) were extracted from the full BNT. The range of scores is 0–15 points.

In Category Fluency (Strauss et al., 2006), participants are given one minute to produce as many different animal names as possible. As it may be challenging to perform fast-paced simultaneous translation of animal names in interpreter-mediated assessments, in this study interpreters were instead instructed to say “yes” for every new animal name and the neuropsychologist put a checkmark for each “yes” on the record form. Immediately following the test, the neuropsychologist checked with the interpreter for any repetitions of animal names. The score is the number of different animal names produced in one minute.

Also, scores from the MMSE and RUDAS were treated as a single measure of general cognitive function (MMSE/RUDAS) in all comparisons. The rationale behind this was that the two instruments have the same range of scores (0–30 points), are highly correlated (Naqvi et al., 2015), have similar diagnostic performance for dementia (Nielsen & Jørgensen, 2020), and were used interchangeably with patients and cognitively intact participants, depending on participant characteristics. In total, 150 participants were assessed with the MMSE and 36 with the RUDAS.

The study adhered to the Declaration of Helsinki for experiments involving humans (reference no. 22007675) and was approved by the Danish Data Protection Agency (RH-2018-34).

### Development of the copenhagen cross-linguistic naming test

The C-CLNT was based on MULTIMAP, a free open-access database of 218 standardized color drawings representing both objects and actions (Gisbert-Muñoz et al., 2021). MULTIMAP includes relevant linguistic variables, including name agreement, frequency (per one million), and number of letters, across several languages (i.e., Spanish, Basque, Catalan, Italian, French, English, German, Mandarin Chinese, and Arabic). However, data on number letters were not included in the present study as this variable seems less relevant when performance time is not an issue. Also, some languages, including Chinese Mandarin, are non-alphabetical languages making this variable inconvenient. MULTIMAP name agreement was established through an online survey with 99 (English) to 128 (Mandarin Chinese) speakers of each language, and frequency data for the words in each language was extracted from text corpora in various online databases (Gisbert-Muñoz et al., 2021).

Based on the original set of MULTIMAP drawings and the procedures described for developing cross-language combinations (Gisbert-Muñoz et al., 2021), an initial set of 38 items (26 objects and 12 actions) was selected by considering MULTIMAP name agreement ( $\geq 80\%$  for objects,  $\geq 75\%$  for actions) data across

**Table 1.** List of items comprising the C-CLNT, and item scores by group

	Item	Control (n = 43)	Affective disorder (n = 20)	MCI (n = 67)	Dementia (n = 56)	P value dementia vs other groups
	<i>Objects</i>					
1	Back	1.00	1.00	.97	.93	.05
2	Sun	1.00	1.00	1.00	1.00	N/A
3	Window	1.00	1.00	1.00	.91	<.001
4	Tree	1.00	1.00	1.00	1.00	N/A
5	Chair	1.00	1.00	1.00	1.00	N/A
6	Ear	1.00	1.00	1.00	.95	.01
7	Mirror	.93	.95	1.00	.91	.04
8	Hat	.98	1.00	1.00	.98	.17
9	Bone	.98	1.00	.85	.56	<.001
10	Lock	.98	1.00	1.00	.93	.01
11	Bear	.98	.95	.95	.89	.03
12	Apple	1.00	1.00	.98	.96	.17
13	Skirt	1.00	.90	.88	.84	.09
14	Fly	.65	.70	.67	.42	.002
15	Umbrella	1.00	1.00	1.00	.98	.03
16	Nail	.88	.85	.98	.76	.002
17	Butterfly	1.00	1.00	.98	.89	<.001
18	Glove	1.00	1.00	1.00	.96	.01
19	Squirrel	.86	.85	.97	.80	.01
20	Scissors	1.00	1.00	1.00	.98	.13
	<i>Actions</i>					
21	Play	1.00	1.00	.97	.93	.05
22	Write	1.00	1.00	1.00	.98	.13
23	Drink	1.00	1.00	1.00	1.00	N/A
24	Eat	1.00	1.00	1.00	.98	.13
25	Push	1.00	1.00	1.00	.96	.03
26	Sing	1.00	1.00	1.00	.96	.03
27	Paint	1.00	1.00	.97	.89	.001
28	Hang	.93	1.00	.97	.85	.004
29	Measure	1.00	1.00	1.00	.98	.13
30	Fish	1.00	.95	1.00	.98	.54

MCI = mild cognitive impairment; N/A = not applicable.

Spanish, Italian, French, English, German, Mandarin Chinese, and Arabic languages. The adopted name agreement cutoff for object items followed the recommendations for developing bilingual naming tests based on the MULTIMAP drawings (Gisbert-Muñoz et al., 2021). However, as action word meanings are more variable across languages than object name meanings (Gentner, 2006), a slightly lower cutoff was used for action items in order to be able to include more items. Subsequently, eight items (*glasses, horse, onion, egg, hand, table, weigh, hunt*) were excluded due to their ambiguity in the Danish cultural context and to reduce cross-language differences in name agreement and frequency.

The final C-CLNT consisted of 30 standardized color drawings (20 objects and 10 actions) with comparable name agreement ( $F(6, 174) = 2.05, p = .06$ ) and frequency  $F(6, 174) = 1.62, p = .11$ ) across seven languages (Supplementary Table S1). The items were ordered according to mean frequency across the target languages and pilot tested in 10 memory clinic patients (2 AD, 4 MCI, 4 affective disorders; 6 male/4 female; mean age  $71.2 \pm 10.8$  years; mean education  $14.3 \pm 2.3$  years). Based on pilot test performances, ambiguity in the scoring criteria for the item *bone* was resolved (i.e., “meat bone” and “chicken bone” was not accepted as correct in Danish). The final set of items selected for the C-CLNT is presented in Table 1, and examples of items are provided in Figure 1.

### Administration and scoring

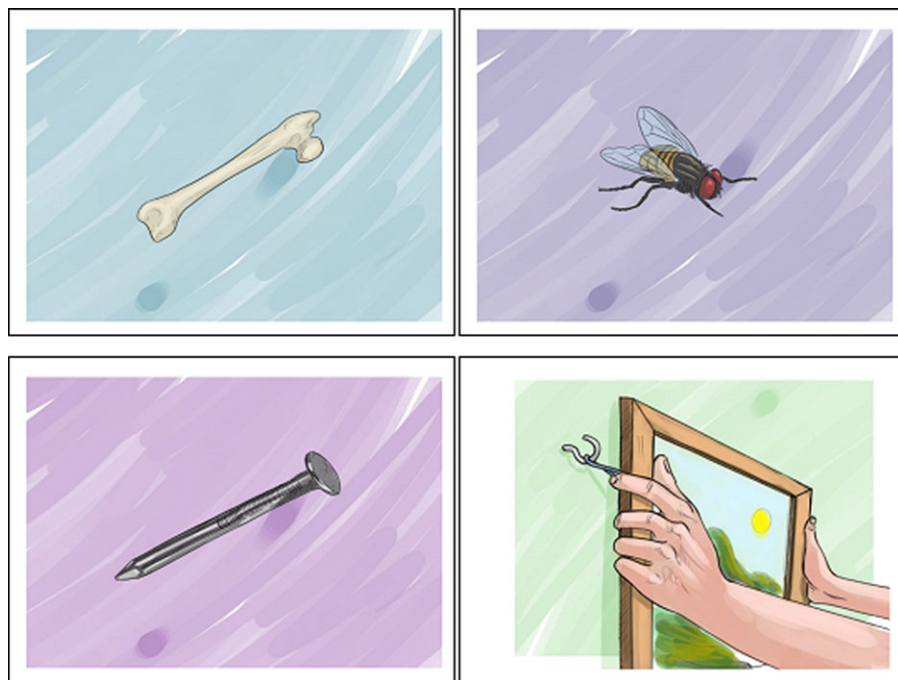
Administration and scoring procedures for the C-CLNT are similar to those of the BNT (Kaplan et al., 2001). Participants are shown each item one at a time and allowed 20 seconds to respond.

When appropriate (e.g., in case of a visual misperception), a semantic cue can be provided. However, a semantic cue is not provided if the incorrect response falls within the same semantic category as the correct response (i.e., if a *nail* is named a “screw” or a *fly* is named a “bee” or “wasp”). If the semantic cue fails to elicit a correct response, a phonetic cue may be given. Participants are allowed 5 seconds to respond following a semantic cue or phonetic cue. There is no discontinuation rule. The administration time is generally < 5 minutes.

The C-CLNT total score is the number of correct responses, including responses after semantic cues. Responses after phonemic cues are not added to the total correct but may be noted to provide qualitative information about naming performance. In the context of multilingualism and inherent language mixing, participants are allowed to respond in any language. A correct response in any language is considered correct.

### Statistical analyses

The significance of group differences on continuous variables was determined using analysis of variance (ANOVA) with pretesting for homogeneity of variances. Welch’s ANOVA was used when the assumption of homogeneity of variances was not met. Effect sizes were calculated as partial eta squared (PES). Fischer’s Exact Test or Pearson’s  $\chi^2$ -test was used to test the significance of group differences in the distribution of categorical variables. Internal consistency of the C-CLNT was determined by coefficient  $\alpha$  as an approximation of scale reliability. To assess construct validity, Spearman’s rank correlation coefficient was used to assess associations between the C-CLNT and traditional language tests.



**Figure 1.** Examples of MULTIMAP items included in the C-CLNT (*bone, fly, nail, hang*). Items presented with permission from the authors.

The effect of years of education, age, sex, and immigrant background on neuropsychological test scores was evaluated using hierarchical regression analyses with plots of residuals as model control. To assess discriminant validity, a receiver operating characteristic curve (ROC) was applied to examine the areas under the curve (AUC), sensitivity, and specificity of the C-CLNT and other language tests for dementia. AUCs were compared using the method proposed by DeLong et al. (1988). Optimal cutoff values were established with Youden's J (calculated as:  $J = \text{sensitivity} + \text{specificity} - 1$ ). All analyses were performed with SPSS version 28.0. A  $p$ -value  $< .05$  (two-tailed) was considered significant.

## Results

A total of 186 participants were included in the study, of which 126 (68%) had a majority ethnic Danish background and 60 (32%) had an immigrant background. Among participants with immigrant backgrounds, 24 originated from a Middle Eastern country, 15 from a South or East Asian country, 13 from another European country, 4 from a North African country, 3 from a Sub-Saharan African country, and one from an Oceanian country. In total, 45 (75%) of the participants with immigrant background had a non-European language as their mother tongue. The mean KSI cultural distance between the original cultures and Danish culture was  $86.2 \pm 19.3$ , ranging from 18.3 (Sweden) to 120.9 (Iraq). Compared to majority ethnic Danish participants, participants with immigrant background were significantly younger ( $68.3$  (range: 42–87) vs  $73.4$  (range: 48–91) years;  $F(1, 183) = 13.93, p < .001$ ) and had fewer years of education ( $10.4$  (range: 0–17) vs  $13.2$  (range: 7–17) years;  $Welch's F(1, 72.73) = 11.89, p < .001$ ). There was no significant difference in sex distribution.

Among patients, 56 were diagnosed with dementia (19 AD, 14 vascular dementia (VaD), 4 mixed AD/VaD, 6 dementia with Lewy bodies/Parkinson's disease dementia, 2 frontotemporal dementia, 6 other specified dementia (normal pressure hydrocephalus, encephalitis, Wernicke-Korsakoff syndrome, HIV-associated

neurocognitive disorder), and 5 unspecified dementia), 67 with MCI, 20 with affective disorder, and 19 with SCD. As patients with SCD did not have formal impairment on neuropsychological testing, or any other neurological or psychiatric diagnosis explaining cognitive complaints (Jessen et al., 2014), they were grouped with cognitively intact participants to form a control group. Participants' characteristics and neuropsychological test performance for the resulting four groups are presented in Table 2.

There were no significant group differences in sex and years of education, but there were significant differences in age ( $F(3, 181) = 10.87, p < .001$ ), proportion of participants with immigrant background ( $\chi^2(3, n = 185) = 15.18, p = .002$ ), and MMSE/RUDAS scores ( $Welch's F(3, 67.74) = 9.36, p < .001$ ). Comparison of performances on the C-CLNT and other language tests displayed significant group differences for the C-CLNT ( $Welch's F(3, 72.73) = 11.74, p < .001, PES = .26$ ), full BNT ( $Welch's F(3, 63.40) = 7.22, p < .001, PES = .09$ ), BNT-15 ( $Welch's F(3, 59.91) = 6.88, p < .001, PES = .08$ ), and Category Fluency ( $Welch's F(3, 64.44) = 26.66, p < .001, PES = .29$ ). Participants with immigrant background obtained significantly lower scores on all language tests compared to participants with majority ethnic Danish background (see Fig. 2). However, differences were considerably lower for the C-CLNT.

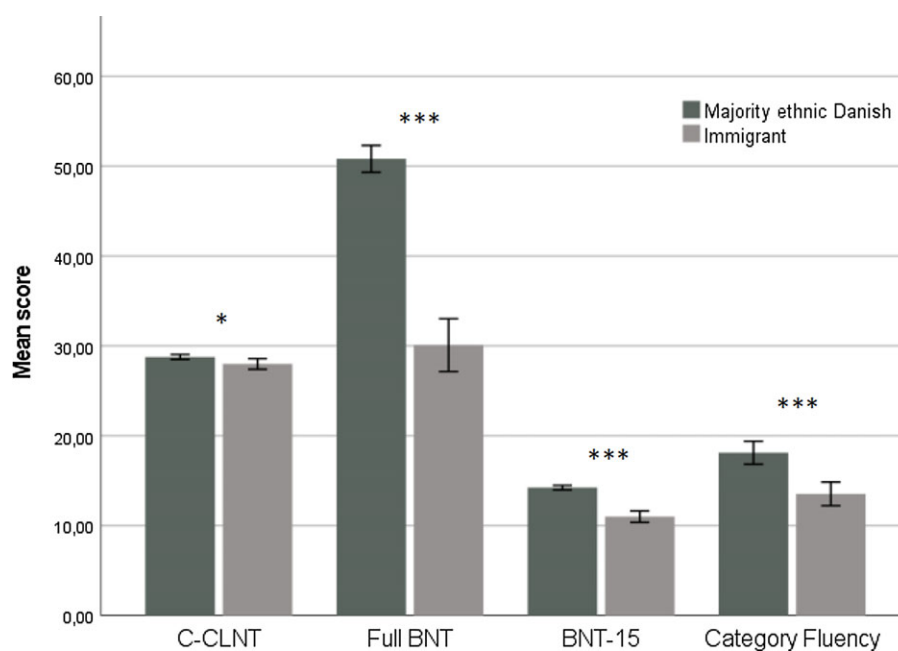
Table 1 lists the items of the C-CLNT ordered according to word frequency. There were four items (i.e., *sun, tree, chair, drink*) showing 100% correct responses across diagnostic groups. Several other items (i.e., *hat, apple, umbrella, scissors, write, eat, measure*) also had very few incorrect responses in all groups. The lowest percentage of correct responses was found on items 9 (*bone*), 13 (*skirt*), 14 (*fly*), 16 (*nail*), and 19 (*squirrel*). Examination of these low-accuracy items revealed typical error types. Participants frequently responded with "meat bone" ("kødben") or "chicken bone" ("kyllingebein") for *bone* (Danish language differentiates between bones in living creatures ("knogle") and bones for consumption), "bee" or "wasp" for *fly*, "screw" for *nail*, "dress" for *skirt* (only males), and "mouse," "rat," or "rabbit" for *squirrel*

**Table 2.** Participant characteristics and neuropsychological test performance

Characteristic	Control ( <i>n</i> = 43)		Affective disorder ( <i>n</i> = 20)		MCI ( <i>n</i> = 67)		Dementia ( <i>n</i> = 56)		<i>p</i> -value
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	
Age	70.7	8.65	62.3	8.05	73.5	8.69	74.0	8.25	<.001
Sex (male/female)	22/20		8/12		37/30		36/19		.20
Years of education	10.9	5.40	12.4	4.14	13.2	3.71	12.3	3.81	.11
Immigrant background ( <i>n</i> )	21		11		13		15		.001
Interpreter used <sup>a</sup> ( <i>n</i> )	14		4		4		7		.11
MMSE/RUDAS	27.8	1.89	27.7	2.37	27.1	2.31	24.9	3.58	<.001
C-CLNT	29.1	1.05	29.1	.85	29.1	.93	27.2	2.26	<.001
Full BNT	42.5	16.22	41.6	14.27	49.3	8.83	41.04	13.01	<.001
BNT-15	13.0	2.27	13.0	2.27	14.0	1.13	12.8	2.45	<.001
Category Fluency	21.4	8.06	17.7	5.59	17.6	5.58	11.6	.29	<.001

BNT = Boston Naming Test; C-CLNT = Copenhagen Cross-Linguistic Naming Test; MCI = mild cognitive impairment; MMSE = mini-mental state examination; RUDAS = Rowland Universal Dementia Assessment Scale.

<sup>a</sup>Comparison based only on participants with an immigrant background.



**Figure 2.** Comparison of test performance in participants with majority ethnic Danish and immigrant backgrounds. Error bars represent 95% CI. \*  $p < .05$ ; \*\*\*  $p < .001$ . BNT = Boston Naming Test; C-CLNT = Copenhagen Cross-Linguistic Naming Test.

(mainly participants with immigrant background). These typical error types constituted 97%, 91%, 74%, 41%, and 26% of all errors on the *bone*, *fly*, *nail*, *skirt*, and *squirrel* item, respectively, and were more frequent in the dementia group compared to the other groups but only significantly so for *bone* (39% vs 19%; Fisher's Exact Test,  $p < .001$ ) and *skirt* (14% vs 5%; Fisher's Exact Test,  $p = .03$ ).

### Scale reliability

Across all participant groups, coefficient  $\alpha$  for C-CLNT was .67 indicating acceptable scale reliability.

### Construct validity

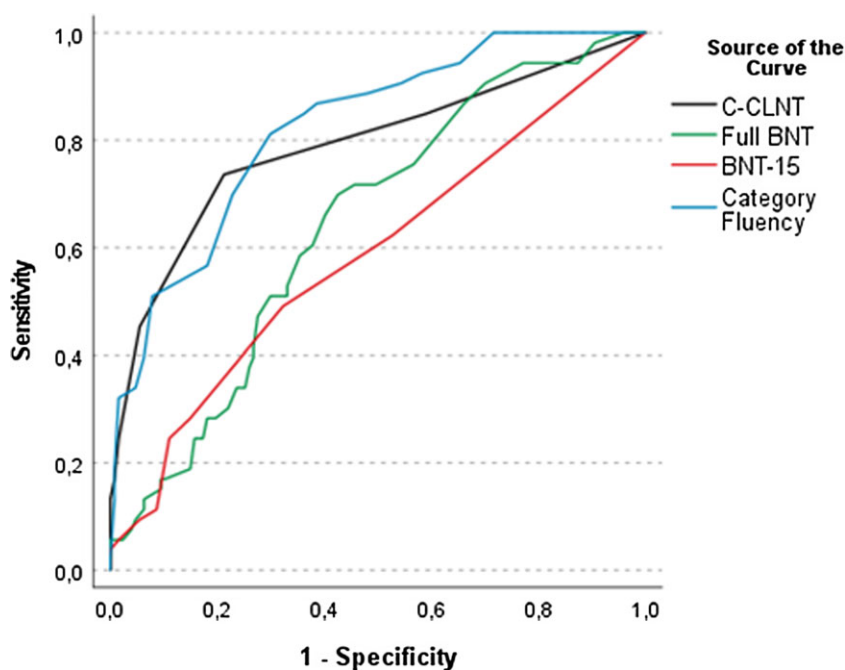
The C-CLNT was significantly related to other language function and general cognitive function tests. The C-CLNT was most robustly correlated with the full BNT ( $r = .55$ ,  $p < .001$ ) and BNT-15 ( $r = .54$ ,  $p < .001$ ), followed by Category Fluency ( $r = .44$ ,  $p < .001$ ), and MMSE/RUDAS ( $r = .35$ ,  $p < .001$ ). Repeating the analyses in a majority of ethnic Danish participants only, correlations with the full BNT ( $r = .72$ ,  $p < .001$ ) and BNT-15

( $r = .67$ ,  $p < .001$ ) were stronger, while the correlation with Category Fluency ( $r = .42$ ,  $p < .001$ ) was similar, and correlation with the MMSE ( $r = .19$ ,  $p = .03$ ) was weaker.

### Discriminant validity

ROC curve analysis revealed that the C-CLNT was highly accurate in discriminating the group of patients with dementia from the other groups (control, affective disorder, MCI). AUCs for the C-CLNT, full BNT, BNT-15, and Category Fluency are illustrated in Figure 3, and AUC values, optimal cutoff scores, sensitivity, and specificity are presented in Table 3. The AUC value for the C-CLNT (AUC = .80) was significantly higher than for the full BNT (AUC = .64,  $z = 3.50$ ,  $p < .001$ ) and BNT-15 (AUC = .59,  $z = 4.46$ ,  $p < .001$ ), but comparable to Category Fluency (AUC = .83). In a subsample of patients with MCI and controls alone ( $n = 109$ ), the AUC for the C-CLNT was .53.

Overall, the accuracy of the C-CLNT in discriminating patients with dementia from the other groups did not significantly differ between participants with majority ethnic Danish and immigrant



**Figure 3.** Receiver operating characteristics curves of C-CLNT, BNT, and Category Fluency for dementia. BNT = Boston Naming Test; C-CLNT = Copenhagen Cross-Linguistic Naming Test.

background (AUC of .80, 95% CI [.70–.88] vs .86, 95% CI [.74–.98],  $z = .83$ ,  $p = .41$ ).

#### Effects of demographic variables

When combining the four groups and correcting for MMSE/RUDAS score, there was a significant positive correlation between years of education and the C-CLNT ( $r = .18$ ,  $p = .01$ ), the full BNT ( $r = .47$ ,  $p < .001$ ), BNT-15 ( $r = .42$ ,  $p < .001$ ), and Category Fluency ( $r = .15$ ,  $p = .05$ ), and between age and the full BNT ( $r = .20$ ,  $p = .008$ ) and BNT-15 ( $r = .21$ ,  $p = .005$ ). Sex was not significantly related to any of the tests.

When the influence of demographic variables on the C-CLNT and other language tests was evaluated with a series of hierarchical regression analyses controlling for MMSE/RUDAS score, significant effects of age and years of education, and immigrant background were present on all tests. However, the variance in test scores explained by immigrant background was 3% for the C-CLNT compared to 8, 28%, and 34% for Category Fluency, BNT-15, and full BNT, respectively (see Table 5).

Repeating the regression analysis in participants with immigrant background and entering years of residence in Denmark, non-European mother tongue, and KSI cultural distance in the last block instead of immigrant background led to a similar picture (see Supplementary Table S2). In these analyses, the variance in test scores explained by years of residence in Denmark, non-European mother tongue, and KSI cultural distance was 3, 8, 18, and 24% for the C-CLNT, Category Fluency, BNT-15, and full BNT, respectively. Adding the use of an interpreter to the regression analyses did not show any significant effects.

#### Abbreviated 20-item version of C-CLNT

An abbreviated version of the C-CLNT was created by excluding the 10 action items included in the full C-CLNT, leaving only the 20 object items. The 20-item C-CLNT was highly correlated with

the full C-CLNT ( $r = .87$ ,  $p < .001$ ) and had comparable psychometric properties. The AUC of the 20-item C-CLNT for dementia was .78, 95% CI [.70–.86], which did not significantly differ from the AUC of full C-CLNT ( $z = 1.21$ ,  $p = .23$ ). At the cutoff  $\leq 18$ , the 20-item C-CLNT had a sensitivity of .83 and a specificity of .70.

#### Discussion

In this study, we described the development and validation of the C-CLNT for the assessment of naming impairment in a culturally, linguistically, and educationally diverse memory clinic patient population in Denmark. The C-CLNT is based on a set of standardized color drawings with data on linguistic variables available across several languages. Items for the final 30-item version of the C-CLNT were selected by considering name agreement and frequency across five European and two non-European languages. The C-CLNT was found to have promising psychometric properties and diagnostic accuracy for detecting naming impairment in culturally, linguistically, and educationally diverse patients with dementia but not MCI. Concerning psychometric properties, the internal reliability of the C-CLNT was acceptable according to standard criteria (coefficient alpha = .67). The convergent validity of the C-CLNT was also good, with scores being moderately to strongly correlated with traditional confrontation naming tests, moderately with Category Fluency, and only weakly with general cognitive functioning. Correlations between the C-CLNT and BNT were strongest in the subsample of majority ethnic Danish participants, reflecting the suboptimal utility of the BNT in culturally, linguistically, and educationally diverse populations.

In the context of the cultural, linguistic, and educational diversity among patients in European memory clinics, it is desirable to have a single standardized confrontation naming test for the cross-linguistic assessment of naming impairment in neurocognitive disorders. The results from this study indicate that the C-CLNT is suitable for detecting naming impairment in

**Table 3.** Diagnostic accuracy of the C-CLNT, BNT, and Category Fluency

	AUC (95% CI)	Optimal cutoff <sup>a</sup>	Sensitivity	Specificity
C-CLNT	.80 (.72–.88)	≤28	.75	.79
Full BNT	.64 (.56–.72)	≤48	.70	.57
BNT-15	.59 (.50–.68)	≤13	.50	.67
Category Fluency	.83 (.77–.89)	≤14	.82	.70

AUC = area under the curve; BNT = Boston Naming Test; CI = confidence interval; C-CLNT = Copenhagen Cross-Linguistic Naming Test.

<sup>a</sup>Optimal cutoff for discriminating between patients with dementia and other groups based on Youden's J.

**Table 4.** Hierarchical regression analysis ( $n = 186$ )

Test	Variables entered	$R^2$	$R^2$ Change	$F$ Change	p-value
C-CLNT	MMSE/RUDAS	.13	.13	26.03	< .001
	Age, years of education	.18	.06	6.07	.003
	Immigrant background	.21	.03	6.56	.01
Full BNT	MMSE/RUDAS	.08	.08	14.96	< .001
	Age, years of education	.30	.22	27.99	< .001
	Immigrant background	.64	.34	168.42	< .001
BNT-15	MMSE/RUDAS	.04	.04	7.84	.01
	Age, years of education	.22	.19	22.04	< .001
	Immigrant background	.49	.28	98.86	< .001
Category Fluency	MMSE/RUDAS	.14	.14	29.28	< .001
	Age, years of education	.12	.02	2.17	.12
	Immigrant background	.24	.08	17.38	< .001

BNT = Boston Naming Test; C-CLNT = Copenhagen Cross-Linguistic Naming Test; MMSE = mini-mental state examination; RUDAS = Rowland Universal Dementia Assessment Scale.

patients referred for neuropsychological evaluation in such a setting. Compared with traditional language tests, the diagnostic accuracy of the C-CLNT for dementia (AUC of .80) was significantly better than that of the full BNT (AUC of .64) and BNT-15 (AUC of .59), but comparable to Category Fluency (AUC of .83). At the cutoff  $\leq 28$ , the sensitivity of the C-CLNT was .75, which indicates that it does not suffer from low sensitivity and diagnostic accuracy as described for several other cross-linguistic naming tests (Abou-Mrad et al., 2017; Araujo et al., 2020; Ardila, 2007; Dick et al., 2002; Gálvez-Lara et al., 2015; Nielsen et al., 2019b). Overall, the diagnostic accuracy of the C-CLNT was slightly lower than that reported for the MINT (AUC of .85; Stassenko et al., 2019), ICMR-PNT (AUC of .81–1.00; Paplikar et al., 2022), and NAME (AUC of .88; Franzen et al., 2023). However, a direct head-to-head comparison of the tests is impossible as study methods and samples differed between studies. In future research, it would be interesting to make a head-to-head comparison of cross-linguistic naming tests in the same study population. Like other (cross-linguistic) naming tests (Li et al., 2022; Paplikar et al., 2022; Stassenko et al., 2019), the C-CLNT showed poor diagnostic accuracy for MCI. This is most likely due to anomia being an uncharacteristic feature of MCI but typically presents in later stages of AD and other dementia disorders (McKhann et al., 2011).

Examination of error types revealed that typical errors generally reflected conceptual deficits and stimulus-bound responses (Rouleau et al., 1992). For instance, participants frequently responded “bee” or “wasp” for fly which was considered to reflect a stimulus-bound response as the drawing used to depict a fly has yellow and black stripes on its lower back (see Fig. 1). Also, patients with dementia more frequently responded “dress” for skirt and “meat bone” or “chicken bone” for bone, which was considered to represent a conceptual deficit as Danish language differentiates between bones in living creatures and bones for consumption, like the difference between cow and beef or pig and pork in English. These error types may not be related to impaired lexical retrieval and anomia but rather reflect impaired semantic memory and/or visual perception, which is also commonly impaired in dementia disorders and known to be important for confrontation naming ability (Taler & Phillips, 2008).

C-CLNT scores were not associated with sex or age, and only negligibly with years of education. Only 6% of the variance in C-CLNT scores was explained by age and years of education, with slightly lower variance explained by immigrant background (3%). In comparison, 34, 28, and 8% of the variance in test scores was explained by immigrant background on the full BNT, BNT-15, and Category Fluency, respectively. Combined with the findings on diagnostic accuracy, these findings support the cross-linguistic properties of the C-CLNT in a diverse memory clinic setting and highlight important limitations of traditional confrontation naming tests. Conversely, when adapting administration procedures to reduce bias in interpreter-mediated assessment, Category Fluency proved to have high clinical utility for cross-linguistic assessment. This is in line with previous reports of Category Fluency being relatively uninfluenced by culture and language (Nielsen et al., 2018), and having high diagnostic accuracy for dementia in multicultural populations (Nielsen et al., 2019b).

An abbreviated 20-item version of the C-CLNT, using only the object items, was highly correlated with the full C-CLNT, which examined both action and object naming, and showed comparable psychometric properties and diagnostic accuracy for dementia. Although assessment of action naming has been suggested to contribute to differential diagnostics as object and action naming may be differentially affected across dementia disorders (Cotelli et al., 2006; Parris & Weekes, 2001), in the present study action naming did not contribute to the overall classification of dementia. Thus, the 20-item C-CLNT may generally be adequate for cross-linguistic assessment of naming impairment in a memory clinic setting. However, future studies might test the effects of action naming with a larger sample of items than used in the present study and a more diverse patient group.

This study has some limitations. Although we were able to analyze C-CLNT performance across patients with affective disorders, MCI, and dementia, the clinical groups were not fully matched on age and proportion of participants with immigrant background, which may have exacerbated some group differences. Furthermore, our dementia sample was too small to analyze the C-CLNT across specific dementia subtypes. Also, the C-CLNT demonstrated a ceiling effect in all clinical groups, except the dementia group, and was not able to discriminate between participants in the control and MCI group. Although the C-CLNT appears to be sensitive to naming impairment in patients with mild dementia, this may indicate that the C-CLNT is not sensitive to more subtle naming impairment. The scale reliability of the C-CLNT was acceptable, but not high. This means that it may not be consistent in measuring naming ability, and the results should



be interpreted with caution. Additionally, in interpreter-mediated assessments, interpreters often struggled with translations of responses for items of the BNT as they did not know the corresponding words in Danish. Also, interpreters assisted in determining whether a nonstandard response was a correct synonym or an incorrect response. On Category Fluency, instructing the interpreters to simply say “yes” for every new animal name did not allow for more careful inspection of repetitions, intrusions, or questionable responses (e.g., same animal name in two languages). Other approaches include having the interpreters write down the responses in the language of their choice or recording and transcribing the responses. However, as the use of interpreters did not differ between clinical groups, these issues are unlikely to have significantly influenced the results. Finally, although the diagnostic accuracy of the C-CLNT did not significantly differ between participants with majority ethnic Danish and immigrant backgrounds, further studies comparing larger cultural and language groups are needed to support the cross-linguistic, cross-cultural, and diagnostic properties of the C-CLNT. Also, reliability metrics, including test-retest, intra-rater, and inter-rater reliability, should be established to provide further support for the psychometric properties of the C-CLNT. As suggested by the European Consortium on Cross-Cultural Neuropsychology (ECCroN) (Franzen et al., 2022), such studies should preferably take several diversity-related variables into account, including limited education and literacy, quality of education, and acculturation.

In conclusion, the novel C-CLNT has promising clinical utility for cross-linguistic assessment of naming impairment in culturally, linguistically, and educationally diverse older adults. Although the C-CLNT was developed by taking into consideration the cultural and linguistic diversity in Europe and was validated in a diverse memory clinic population, the C-CLNT may also be suitable for assessment of naming impairment in other cultural and clinical contexts, including culturally and linguistically diverse populations in other world regions, and patients with stroke or traumatic brain injury. However, before such applications further research is needed to establish the utility of the C-CLNT in these contexts.

**Supplementary material.** The supplementary material for this article can be found at <https://doi.org/10.1017/S1355617723000437>.

**Acknowledgments.** The authors report no competing interests. This research was supported by THE VELUX FOUNDATIONS (grant number 00042578), which had no role in the formulation of research questions, choice of study design, data collection, data analysis, or decision to publish. The Danish Dementia Research Centre is supported by the Danish Ministry of Health. The authors would like to thank all neuropsychologists at the Copenhagen University Hospital Memory Clinic at Rigshospitalet who contributed to the data collection, and Kasper Jørgensen who contributed with constructive comments to the article draft.

## References

- Abou-Mrad, F., Chelune, G., Zamrini, E., Tarabey, L., Hayek, M., & Fadel, P. (2017). Screening for dementia in Arabic: Normative data from an elderly Lebanese sample. *The Clinical Neuropsychologist*, 31(sup1), 1–19.
- Araujo, N. B., Nielsen, T. R., Barca, M. L., Engedal, K., Marinho, V., Deslandes, A. C., Coutinho, E. S., & Laks, J. (2020). Brazilian version of the European cross-cultural neuropsychological test battery (CNTB-BR): Diagnostic accuracy across schooling levels. *Brazilian Journal of Psychiatry*, 42(3), 286–294.
- Ardila, A. (2007). Toward the development of a cross-linguistic naming test. *Archives of Clinical Neuropsychology*, 22(3), 297–307.

- American Psychological Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC: American Psychiatric Publishing.
- Baird, A. D., Ford, M., & Podell, K. (2007). Ethnic differences in functional and neuropsychological test performance in older adults. *Archives of Clinical Neuropsychology*, 22(3), 309–318.
- Bertola, L., & Malloy-Diniz, L. F. (2018). Assessing knowledge: Psychometric properties of the BAMS semantic memory battery. *Archives of Clinical Psychiatry (São Paulo)*, 45(2), 33–37.
- Boone, K. B., Victor, T. L., Wen, J., Razani, J., & Ponton, M. (2007). The association between neuropsychological scores and ethnicity, language, and acculturation variables in a large patient population. *Archives of Clinical Neuropsychology*, 22(3), 355–365. [https://doi.org/S0887-6177\(07\)00017-0](https://doi.org/S0887-6177(07)00017-0).
- Cotelli, M., Borroni, B., Manenti, R., Alberici, A., Calabria, M., Agosti, C., Arévalo, A., Ginex, V., Ortelli, P., Binetti, G. (2006). Action and object naming in frontotemporal dementia, progressive supranuclear palsy, and corticobasal degeneration. *Neuropsychology*, 20(5), 558–565.
- DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more correlated receiver operating characteristic curves: A nonparametric approach. *Biometrics*, 44(3), 837–845.
- Dick, M. B., L.Teng, E., Kempler, D., S.Davis, D., & Taussig, I. M. (2002). The cross-cultural neuropsychological test battery (CNNB): Effects of age, education, and cognitive status on performance. In R. Ferraro (Eds.). *Minority and Cross-Cultural Aspects of Neuropsychological assessment*. (pp. 15–39). Swets & Zeitlinger.
- Folstein, M. F., Folstein, S. E., & McHugh, P. R. (1975). “Mini-mental state”: A practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3), 189–198.
- Franzen, S., Neuropsychology, E. C. C.-C., Watermeyer, T. J., Pomati, S., Papma, J. M., Nielsen, T. R., Narme, P., Mukadam, N., Lozano-Ruiz, Á., Ibanez-Casas, I. (2022). Cross-cultural neuropsychological assessment in Europe: Position statement of the European consortium on cross-cultural neuropsychology (ECCroN). *The Clinical Neuropsychologist*, 36(3), 546–557.
- Franzen, S., Papma, J. M., van den Berg, E., & Nielsen, T. R. (2020). Cross-cultural neuropsychological assessment in the European union: A Delphi expert study. *Archives of Clinical Neuropsychology*, 36(5), 815–830.
- Franzen, S., van den Berg, E., Ayhan, Y., Satoer, D., Türkoğlu, Ö., Akpulat, G. G., Visch-Brink, E., Scheffers, E., Kranenburg, J., Jiskoot, L. (2023). The naming assessment in multicultural Europe (NAME): Development and validation in a multicultural memory clinic. *Journal of the International Neuropsychological Society*, 29(1), 92–104.
- Gálvez-Lara, M., Moriana, J. A., Vilar-López, R., Fasfous, A. F., Hidalgo-Ruzzante, N., & Pérez-García, M. (2015). Validation of the cross-linguistic naming test: a naming test for different cultures? a preliminary study in the Spanish population. *Journal of Clinical and Experimental Neuropsychology*, 37(1), 102–112.
- Gentner, D. (2006). Why verbs are hard to learn. In K. Hirsh-Pasek, & R. Golinkoff (Eds.). *Action meets word: How children learn verbs*. Oxford University Press.
- George, A., & Mathuranath, P. (2007). Community-based naming agreement, familiarity, image agreement and visual complexity ratings among adult Indians. *Annals of Indian Academy of Neurology*, 10(2), 92.
- Gisbert-Muñoz, S., Quiñones, I., Amoroso, L., Timofeeva, P., Geng, S., Boudelaa, S., Pomposo, I., Gil-Robles, S., & Carreiras, M. (2021). MULTIMAP: Multilingual picture naming test for mapping eloquent areas during awake surgeries. *Behavior Research Methods*, 53(2), 918–927.
- Gollan, T. H., Fennema-Notestine, C., Montoya, R. I., & Jernigan, T. L. (2007). The bilingual effect on Boston naming test performance. *Journal of the International Neuropsychological Society*, 13(2), 197–208. <https://doi.org/S135561770700038>.
- Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A multi-lingual naming test (MINT) and preliminary norms for young and aging Spanish-English bilinguals. *Biling (Camb Engl)*, 15(3), 594–615. <https://doi.org/10.1017/S1366728911000332>.
- Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., Ogar, J. M., Rohrer, J. D., Black, S., Boeve, B. F. (2011).

- Classification of primary progressive aphasia and its variants. *Neurology*, 76(11), 1006–1014.
- Grossman, M., McMillan, C., Moore, P., Ding, L., Glosser, G., Work, M., & Gee, J. (2004). What's in a name: Voxel-based morphometric analyses of MRI and naming difficulty in Alzheimer's disease, frontotemporal dementia and corticobasal degeneration. *Brain*, 127(3), 628–649.
- Ivanova, I., Salmon, D. P., & Gollan, T. H. (2013). The multilingual naming test in Alzheimer's disease: Clues to the origin of naming impairments. *Journal of the International Neuropsychological Society*, 19(3), 272–283.
- Ivanova, M. V., & Hallowell, B. (2013). A tutorial on aphasia test development in any language: Key substantive and psychometric considerations. *Aphasiology*, 27(8), 891–920.
- Jessen, F., Amariglio, R. E., Van Boxtel, M., Breteler, M., Ceccaldi, M., Chételat, G., Dubois, B., Dufouil, C., Ellis, K. A., Van Der Flier, W. M. (2014). A conceptual framework for research on subjective cognitive decline in preclinical Alzheimer's disease. *Alzheimer's & Dementia*, 10(6), 844–852.
- Jørgensen, K., Johannsen, P., & Vogel, A. (2017). A Danish adaptation of the Boston naming test: Preliminary norms for older adults and validity in mild Alzheimer's disease. *The Clinical Neuropsychologist*, 31(sup1), 72–87.
- Kaplan, E. F., Goodglass, H., & Weintraub, S. (2001). The Boston Naming Test. In *In* (2nd ed.). Philadelphia, USA: Lippincott Williams & Wilkins.
- Kogut, B., & Singh, H. (1988). The effect of national culture on the choice of entry mode. *Journal of International Business Studies*, 19(3), 411–432.
- 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.
- Li, C., Zeng, X., Neugroschl, J., Aloysi, A., Zhu, C. W., Xu, M., Teresi, J. A., Oceppek-Welikson, K., Ramirez, M., Joseph, A. (2022). The 32-item multilingual naming test: Cultural and linguistic biases in monolingual chinese-speaking older adults. *Journal of the International Neuropsychological Society*, 28(5), 511–519.
- Maruta, C., Guerreiro, M., de, M. A., Hort, J., & Scheltens, P. (2011). The use of neuropsychological tests across Europe: The need for a consensus in the use of assessment tools for dementia. *European Journal of Neurology*, 18(2), 279–285.
- Mathuranath, P. S., Nestor, P. J., Berrios, G. E., Rakowicz, W., & Hodges, J. R. (2000). A brief cognitive test battery to differentiate Alzheimer's disease and frontotemporal dementia. *Neurology*, 55(11), 1613–1620.
- McKeith, I. G., Boeve, B. F., Dickson, D. W., Halliday, G., Taylor, J.-P., Weintraub, D., Aarsland, D., Galvin, J., Attems, J., Ballard, C. G. (2017). Diagnosis and management of dementia with Lewy bodies: Fourth consensus report of the DLB consortium. *Neurology*, 89(1), 88–100.
- McKhann, G. M., Knopman, D. S., Chertkow, H., Hyman, B. T., Jack, C. R., Jr., Kawas, C. H., Klunk, W. E., Koroshetz, W. J., Manly, J. J., Mayeux, R., Mohs, R. C., Morris, J. C., Rossor, M. N., Scheltens, P., Carrillo, M. C., Thies, B., Weintraub, S., Phelps, C. H. (2011). The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & Dementia*, 7(3), 263–269.
- Morris, J. C., Mohs, R. C., Rogers, H., Fillenbaum, G., & Heyman, A. (1988). Consortium to establish a registry for Alzheimer's disease (CERAD) clinical and neuropsychological assessment of Alzheimer's disease. *Psychopharmacology Bulletin*, 24(4), 641–652.
- Naqvi, R. M., Haider, S., Tomlinson, G., & Alibhai, S. (2015). Cognitive assessments in multicultural populations using the rowland universal dementia assessment scale: A systematic review and meta-analysis. *CMAJ*, 187(5), E169–E175. <https://doi.org/cmaj.140802>.
- Nielsen, T. R. (2022). Cognitive assessment in culturally, linguistically, and educationally diverse older populations in Europe. *American Journal of Alzheimer's Disease & Other Dementias*<sup>®</sup>, 37, 15333175221117006.
- Nielsen, T. R., Franzen, S., Goudsmit, M., & Uysal-Bozkir, Ö. (2022). Cross-cultural neuropsychological assessment in the European context: embracing maximum cultural diversity at minimal geographic distances. In *Cultural diversity in neuropsychological assessment*. (pp. 313–325). Routledge.
- Nielsen, T. R., & Jørgensen, K. (2020). Cross-cultural dementia screening using the rowland universal dementia assessment scale: A systematic review and meta-analysis. *International Psychogeriatrics*, 32(9), 1–14. <https://doi.org/10.1017/S1041610220000344>.
- Nielsen, T. R., Segers, K., Vanderaspoilden, V., Beinhoff, U., Minthon, L., Pissioti, A., Bekkhus-Wetterberg, P., Bjorklof, G. H., Tsolaki, M., Gkioka, M., Waldemar, G. (2019a). Validation of a brief multicultural cognitive examination (MCE) for evaluation of dementia. *International Journal of Geriatric Psychiatry*, 34(7), 982–989. <https://doi.org/10.1002/gps.5099>.
- Nielsen, T. R., Segers, K., Vanderaspoilden, V., Beinhoff, U., Minthon, L., Pissioti, A., Bekkhus-Wetterberg, P., Bjorklof, G. H., Tsolaki, M., Gkioka, M., Waldemar, G. (2019b). Validation of a European cross-cultural neuropsychological test battery (CNTB) for evaluation of dementia. *International Journal of Geriatric Psychiatry*, 34(1), 144–152. <https://doi.org/10.1002/gps.5002>.
- Nielsen, T. R., Segers, K., Vanderaspoilden, V., Bekkhus-Wetterberg, P., Minthon, L., Pissioti, A., Bjorklof, G. H., Beinhoff, U., Tsolaki, M., Gkioka, M. (2018). Performance of middle-aged and elderly European minority and majority populations on a cross-cultural neuropsychological test battery (CNTB). *The Clinical Neuropsychologist*, 32(8), 1411–1430.
- Paplikar, A., Varghese, F., Alladi, S., Vandana, V., Darshini, K., Iyer, G. K., Kandukuri, R., Divyaraj, G., Sharma, M., Dhaliwal, R. (2022). Picture-naming test for a linguistically diverse population with cognitive impairment and dementia. *International Journal of Language & Communication Disorders*, 57(4), 881–894.
- Parris, B., & Weekes, B. (2001). Action naming in dementia. *Neurocase*, 7(6), 459–471.
- 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.
- Rascovsky, K., Hodges, J. R., Knopman, D., Mendez, M. F., Kramer, J. H., Neuhaus, J., Van Swieten, J. C., Seelaar, H., Dopper, E. G., Onyike, C. U. (2011). Sensitivity of revised diagnostic criteria for the behavioural variant of frontotemporal dementia. *Brain*, 134(9), 2456–2477.
- Reis, A., Faisca, L., Ingvar, M., & Petersson, K. M. (2006). Color makes a difference: Two-dimensional object naming in literate and illiterate subjects. *Brain and Cognition*, 60(1), 49–54. (Not in File).
- Reis, A., Petersson, K. M., Castro-Caldas, A., & Ingvar, M. (2001). Formal schooling influences two- but not three-dimensional naming skills. *Brain and Cognition*, 47(3), 397–411.
- 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.
- Rouleau, I., Salmon, D. P., Butters, N., Kennedy, C., & McGuire, K. (1992). Quantitative and qualitative analyses of clock drawings in Alzheimer's and Huntington's disease. *Brain and Cognition*, 18(1), 70–87.
- Sachdev, P., Kalaria, R., O'Brien, J., Skoog, I., Alladi, S., Black, S. E., Blacker, D., Blazer, D. G., Chen, C., Chui, H., Ganguli, M., Jellinger, K., Jeste, D. V., Pasquier, F., Paulsen, J., Prins, N., Rockwood, K., Roman, G., Scheltens, P., & Cognitive, D. (2014). Diagnostic criteria for vascular cognitive disorders: A VASCOG statement. *Alzheimer Disease and Associated Disorders*, 28(3), 206–218. <https://doi.org/10.1097/WAD.0000000000000034>.
- Stasenko, A., Jacobs, D. M., Salmon, D. P., & Gollan, T. H. (2019). The multilingual naming test (MINT) as a measure of picture naming ability in Alzheimer's disease. *Journal of the International Neuropsychological Society*, 25(8), 821–833.
- Storey, J. E., Rowland, J. T., Basic, D., Conforti, D. A., & Dickson, H. G. (2004). The rowland universal dementia assessment scale (RUDAS): A multicultural cognitive assessment scale. *International Psychogeriatrics*, 16(1), 13–31.
- Strain, J. F., Didehban, N., Spence, J., Conover, H., Bartz, E. K., Mansinghani, S., Jeroudi, M. K., Rao, N. K., Fields, L. M., Kraut, M. A. (2017). White matter changes and confrontation naming in retired aging national football league athletes. *Journal of neurotrauma*, 34(2), 372–379.
- Strauss, E., Sherman, E. M. S., & Spreen, O. (2006). A compendium of neuropsychological tests. In *Administration, norms, and commentary* (3rd ed.). Oxford University Press.
- Stålhammar, J., Hellström, P., Eckerström, C., & Wallin, A. (2022). Neuropsychological test performance among native and non-native Swedes: Second language effects. *Archives of Clinical Neuropsychology*, 37(4), 826–838.

- Taler, V., & Phillips, N. A. (2008). Language performance in Alzheimer's disease and mild cognitive impairment: A comparative review. *Journal of Clinical and Experimental Neuropsychology*, 30(5), 501–556.
- The REhabilitation and recovery of peopLE with Aphasia after StrokE (RELEASE) Collaborators (2021). The REhabilitation and recovery of peopLE with aphasia after stroke (RELEASE) collaborators, predictors of poststroke aphasia recovery: A systematic review-informed individual participant data meta-analysis. *Stroke*, 52(5), 1778–1787.
- Van Mol, C., & De Valk, H. (2016). Migration and immigrants in Europe: A historical and demographic perspective. In *Integration processes and policies in Europe* (pp. 31–55). Cham: Springer.
- Weeks, S. K., McGann, P. E., Michaels, T. K., & Penninx, B. W. (2003). Comparing various short-form geriatric depression scales leads to the GDS-5/15. *Journal of Nursing Scholarship*, 35(2), 133–137. <https://doi.org/10.1111/j.1547-5069.2003.00133.x>
- Winblad, B., Palmer, K., Kivipelto, M., Jelic, V., Fratiglioni, L., Wahlund, L. O., Nordberg, A., Bäckman, L., Albert, M., Almkvist, O. (2004). Mild cognitive impairment-beyond controversies, towards a consensus: Report of the international working group on mild cognitive impairment. *Journal of Internal Medicine*, 256(3), 240–246.
- World Health Organization. (1993). *The ICD 10 classification of mental and behavioral disorders: Diagnostic criteria for research*. Geneva, Switzerland: World Health Organization.