REVIEW

Neuropsychological Tests of Memory, Visuospatial, and Language Function in Parkinson's Disease: Review, Critique, and Recommendations

Ondrej Bezdicek, PhD,[1*] Roberta Biundo, PhD,[2,3] Sarai Boelema, PhD,[4] Davide Maria Cammisuli, PhD,[5] Brenna Cholerton, PhD,[6] Alice Cronin-Golomb, PhD,[7] John C. Dalrymple-Alford, PhD,[8,9,10] Annelien Duits, PhD,[11] Robert Fellows, PhD,[12] Adam Gerstenecker, PhD,[13] Hanane El Hachioui, PhD,[4] Hana Horáková, PhD,[14] Janneke Koerts, PhD,[15] Bonnie Levin, PhD,[16] Inga Liepelt-Scarfone, PhD,[17,18,19] Marina Sarno, PsyD,[16] Tiago A. Mestre, PhD,[20] Álvaro Sánchez Ferro, MD, PhD,[21,22] Michelle Hyczy de Siqueira Tosin, PhD,[23] Matej Skorvanek, MD, PhD,[24,25] Daniel Weintraub, MD,[26] Gert J. Geurtsen, PhD,[4,27] and
the members of the IPMDS Clinical Outcome Assessment Scientific Evaluation Committee

ABSTRACT: Background: Cognitive impairment in Parkinson's disease (PD) is a key non-motor complication during the disease course.

Objectives: A review of detailed cognitive instruments to detect mild cognitive impairment (PD-MCI) or dementia (PDD) is needed to establish optimal tests that facilitate diagnostic accuracy.

Methods: We performed a systematic literature review of tests that assess memory, language including premorbid intelligence, and visuospatial domains (for tests of attention and executive functions see accompanying review) to determine suitability to assess cognition in PD. Based on in-depth scrutiny of psychometric and other relevant clinimetric properties, tests were rated as "recommended," "recommended with caveats," "suggested," or "listed" by the International Parkinson and Movement Disorder Society (IPMDS) panel of experts according to the IPMDS Clinical Outcome Assessment Scientific Evaluation Committee guidelines.

Results: We included 39 tests encompassing 48 outcome measures. Seven tests (different versions or subtests of the test counted once) were recommended, including four for memory, one for visuospatial domains, one for language (including three measures), and one for estimated premorbid intelligence. Furthermore, 10 tests (12 measures) were "recommended with caveats," 11 were "suggested," and 11 (15 measures) were "listed."

Conclusions: Recommended neuropsychological tests in memory, visuospatial functions, and language are proposed to guide the assessment of cognitive impairment and its progression in PD-MCI and PDD, and for use in clinical trials to stratify participants or as outcome measures. Novel measures being developed will need extensive validation research to be "recommended." © 2025 The Author(s). *Movement Disorders* published by Wiley Periodicals LLC on behalf of International Parkinson and Movement Disorder Society.

Key Words: clinimetric; cognitive; dementia; neuropsychology; Parkinson's disease; rating scales; test

[1]Department of Neurology and Centre of Clinical Neuroscience, First Faculty of Medicine and General University Hospital in Prague, Charles University, Prague, Czech Republic; [2]Department of General Psychology, University of Padua, Padua, Italy; [3]Parkinson and Movement Disorders Unit, Study Center for Neurodegeneration (CESNE), Center for Rare Neurological Diseases (ERN-RND), Department of Neuroscience, University of Padua, Padua, Italy; [4]Amsterdam University Medical Center, University of Amsterdam, Department of Medical Psychology, Amsterdam, The Netherlands; [5]Department of Psychology, Catholic University of the Sacred Heart, Milan, Italy; [6]Department of Pathology, Stanford University School of Medicine, Stanford, California, USA; [7]Department of Psychological and Brain Sciences, Boston University, Boston, Massachusetts, USA; [8]Te Kura Mahi ā-Hirikapo, School of Psychology, Speech and Hearing, University of Canterbury, Christchurch, New Zealand; [9]New Zealand Brain Research Institute, Christchurch, New Zealand; [10]Department of Medicine, University of Otago, Christchurch, New Zealand; [11]Department of Medical Psychology, Maastricht University Medical Centre and Radboud University Centre Nijmegen and the Department of Psychiatry and Neuropsychology, Maastricht University, Maastricht, The Netherlands; [12]Department of Psychiatry, University of California, San Diego, California, USA; [13]Department of Neurology, Division of Neuropsychology, University of Alabama at Birmingham, Birmingham, Alabama, USA; [14]Memory Clinic, Department of Neurology, Second Faculty of Medicine, Charles University and Motol University Hospital, Prague, Czech Republic; [15]Department of Clinical and Developmental Neuropsychology, Faculty of Behavioral and Social Sciences, University of Groningen, Groningen, The Netherlands; [16]Division of Neuropsychology and Cognitive Neuroscience, Department of Neurology, University of Miami Miller School of Medicine, Miami, Florida, USA; [17]Hertie Institute for Clinical Brain Research, Department of Neurodegenerative Diseases, Eberhard Karls University Tübingen, Tübingen, Germany; [18]German Center for Neurodegenerative Diseases (DZNE), Tübingen, Germany; [19]IB-Hochschule, Stuttgart, Germany; [20]Parkinson's Disease and Movement Disorders Center, Division of

Cognitive impairment in Parkinson's disease (PD) is one of the most common and important non-motor, mental health issues,[1] which affect patients' well-being, caregiver burden, healthcare costs, severity of depressive symptoms, and instrumental activities of daily living.[2-7] Cognitive impairment in persons with PD may evolve either early or years after disease onset. It can also develop more insidiously in the preclinical or prodromal phase.[8-12] Cognitive impairment stages are conceptualized as PD with mild cognitive impairment (PD-MCI),[13,14] a pre-dementia state, and as formal dementia in PD (PDD).[14,15] The prevalence of cognitive impairment ranges between 25% and 42.5% for PD-MCI in newly diagnosed PD cases to the majority who survive more than 10 years of disease progression, and perhaps as many as 80% after 20 years from clinical onset of the disease, developing PDD.[9,11,16-19]

A dysexecutive syndrome has often been regarded as a hallmark of the clinical phenotype of cognitive impairment in PD. However, research shows that the cognitive profile is heterogeneous and associated with the disruption of multiple neurotransmitter systems and the spread of Lewy bodies and neurites across many brain structures. There may also be contributions from non-synuclein pathology, including Alzheimer's disease (AD) and cerebrovascular disease. Indeed, deficits have been found in memory, attention, visual perception, visuospatial skills, and language.[20-25] This broader spectrum of impairments led to revisions and validation of new diagnostic criteria for PD-MCI and PDD by the International Parkinson and Movement Disorder Society (IPMDS) study groups.[14,15,26]

Neuropsychological assessment at the more detailed level II requires a comprehensive examination of five cognitive functions (attention and working memory; executive functions; language; memory; and visuospatial functions) with at least two tests per domain.[14] It is an optimal choice for reaching validity and diagnostic accuracy when obtaining a diagnosis of PD-MCI or for predicting the risk for the development of PDD.[13,26-28] However, the most effective tests to evaluate each domain in individuals with PD are unknown.

These tests should be appropriate for use in randomized clinical trials, including stratification of patients at enrolment or as an outcome measure together with trials testing possible neuroprotective therapies.[29,30]

To address this, the IPMDS commissioned a review to evaluate available tests that focus on language, memory, and visuospatial functions (see an accompanying review on executive functions and attention). The psychometric properties of tests[31] were scrutinized as a follow-up of the review of "global tests" for cognitive screening, level I assessment, in PD under the IPMDS Clinical Outcome Assessment (COA) Scientific Evaluation Committee (SEC) guidelines.[32] The selection of tests was based on expert reviews of neuropsychological tests used in PD and their psychometric properties in this population.

Methods

Organization and Review Process

An international group of experts on neuropsychological assessment in PD was selected by the IPMDS COA SEC. O.B. chaired the group focused on reviewing language, memory, and visuospatial function tests. Sixteen experts selected, reviewed and critiqued measures evaluating key aspects of these tests of cognitive function in PD (S.B., D.M.C., B.C., A.C.G., J.C.D.A., A.D., R.F., A.G., H.E.H., H.H., J.K., B.L., I.L.S., M.S., R.B., G.J.G., A.S.F, T.A.M., and M.H.S.T.). Each measure was evaluated using the following systematic procedure: all reviews were entered into a template provided by the IPMDS COA SEC and adapted for neuropsychological test review. Each test review encompassed the description of the test/scale, properties, contemporary use, psychometric properties, and overall evaluation of the suitability and applicability of the test in a clinical setting, especially to PD patients. Each scale/test was

Neurology, Department of Medicine, The Ottawa Hospital Research Institute, University of Ottawa Brain and Mind Institute, Ottawa, Ontario, Canada; [21]Movement Disorders Unit, Neurology Department, Hospital Universitario 12 de Octubre, Madrid, Spain; [22]Centro de Investigacion Biomédica en Red sobre Enfermedades Neurodegenerativas (CIBERNED), Madrid, Spain; [23]Department of Neurological Sciences, Rush University Medical Center, Chicago, Illinois, USA; [24]Department of Neurology, Safarik University, Kosice, Slovakia; [25]Department of Neurology, University Hospital of L. Pasteur, Kosice, Slovakia; [26]Department of Psychiatry, Perelman School of Medicine at the University of Pennsylvania and Parkinson's Disease Research, Education and Clinical Center (PADRECC), Philadelphia Veterans Affairs Medical Center, Philadelphia, Pennsylvania, USA; [27]Amsterdam Neuroscience, Neurodegeneration, Amsterdam, The Netherlands

*Correspondence to: Prof. Ondrej Bezdicek, Department of Neurology and Centre of Clinical Neuroscience, First Faculty of Medicine and General University Hospital in Prague, Charles University, Kateřinská 30, 128 21 Praha 2, Czech Republic; E-mail: ondrej.bezdicek@lf1.cuni.cz

Relevant conflicts of interest/financial disclosures: No specific funding was received for this work. The authors declare that there are no conflicts of interest relevant to this work.

Funding agency: None.

Received: 25 October 2024; **Revised:** 14 February 2025; **Accepted:** 18 February 2025

Published online 27 March 2025 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/mds.30166

independently evaluated first by two neuropsychologists, and subsequently by the chair of the group. If the latter disagreed with the reached level decision, provided by each selected panel member, an intermediate-step discussion about the recommendation level was requested and an overall consensus was achieved. The final decision was based on consensus among all expert panel members. Two liaisons (M.S., D.W.) oversaw and reviewed the overall project. Finally, the manuscript was reviewed and approved by the IPMDS COA SEC chairs (M.S., M.H.S.T.), COA program directors (T.A.M., A.S.F.), and members of the IPMDS COA SEC.

Literature Search

The current review followed guidelines and practices introduced by the previous reviews of IPMDS task forces.[32,33] The literature search was done by using PubMed, Web of Science, Medline, and Scopus for all publications from 1975 to December 2022. Keywords used in the search contained "Parkinson*" and the terms "cognit*" OR "test" OR "neuropsych*" OR "cognition" OR "cognitive deficits" OR "neuropsychological assessment" OR "cognitive testing" OR "neurocognitive" OR "neurocognitive assessment" OR "screening" OR "evaluation." Accepted for the review were already-published or in-press peer-reviewed articles available to the task force members covering language, memory, and visuospatial function tests. Premorbid intelligence tests were evaluated separately.

Selection of Tests

The tests included were those used in PD-MCI diagnostic criteria[14] or those used at least once in PD research and covering specifically memory, language, or visuospatial functioning. Additionally, it could be a single specific measure in a multi-test battery. The most recent versions of tests undergoing re-standardization, unstandardized tests, or (commercially) unavailable measures were excluded. Computerized neuropsychological tests were also excluded because they would not necessarily be available in all clinical settings. The final exclusion criterion was tests whose English version was not available or had copyright issues. The committee decided to include tests of premorbid intelligence separately, which are also treated in this way in tables. Comparison to premorbid level of functioning is a standard for defining the presence and extent of cognitive impairment.[34]

Recommendation Levels

The recommendation criteria were adopted from the previous reviews:[27] "recommended," "recommended with caveats," "suggested," and "listed." Each test measure was categorized as follows. A test was "recommended" if (1) it had been applied to PD populations; and (2) there are data on its use in studies beyond the group that developed the test; and (3) it had been studied clinimetrically in PD and found to be valid, reliable, and sensitive to change. "Recommended with caveats" means the test's properties were generally found to be adequate, but some of the measurement properties were not evaluated specifically at different stages of cognitive impairment in PD. A test was "suggested" if it had been applied to PD populations, but only one of the other criteria applied. A test is "listed" if it was used in the PD population, but did not meet the other two criteria defined for recommended tests.[32] Measures of premorbid intelligence have been designed to be resistant to changes occurring as a result of a psychiatric or neurological disorder.

Results

Identified Tests and Their Use in Clinical Research

Reporting on tests follows (1) clinimetric properties, including (i) reliability (internal consistency, intra-rater, inter-rater and/or test–retest reliability); (ii) validity (including construct and empirical validity indices); (iii) sensitivity to change (from longitudinal studies or clinical trials); (2) strengths and weaknesses; and (3) level of recommendation and justification.[32] The order of and a concise evaluation of reviewed and recommended tests can be found in Table 1. For brevity, only "recommended" tests are summarized below.

Memory	*Language*
• Rey's Auditory Verbal Learning Test (RAVLT)	• Boston Naming Test (BNT-60; BNT-30; BNT-15)
• California and Philadelphia Verbal Learning Test (CVLT/CVLT-II/CVLT3 and PVLT)	
• Hopkins Verbal Learning Test (HVLT/HVLT-R)	
• Rivermead Behavioural Memory Test (RBMT/RBMT II/RBMT III)	

Visuospatial Function	*Estimated Premorbid Intelligence*
• Wechsler Adult Intelligence Scale (WAIS-III/IV: Matrix Reasoning)	• National Adult Reading Test (NART/NAART/NAART-R)

All other reviewed tests can be found in Table S1.

Overall, we identified 39 neuropsychological tests encompassing 48 measures (including subtests or subscales, but not counting as different upgraded versions of the same measure) focused on language, memory, or visuospatial function domains and that fulfilled the inclusion criteria. The review found evidence to "recommend" seven tests, one for language (all Boston Naming

TABLE 1 *Recommended neuropsychological tests including their psychometric properties*

Test	Reliability	Validity	Sensitive to change	Strengths	Clinimetric limitations	Recommendation level
Memory						
RAVLT	Good	Good	Sensitive to progression of memory impairment during PD course	• Multiple normative datasets • Excellent psychometric properties	• Increased difficulty for PDD patients • Unsuitable for screening	Recommended
CVLT/CVLT-II/ CVLT3 and PVLT	Good	Good	Sensitive to progression of memory impairment during PD course	• Multiple normative datasets • Excellent psychometric properties • Includes also cued recall	• Increased difficulty for PDD patients • Unsuitable for screening	Recommended
HVLT/HVLT-R	Adequate	Adequate	Sensitive to progression of memory impairment during PD course	• Multiple normative datasets • Suitable for repeated testing	• Unsuitable for screening	Recommended
RBMT/RBMT II/RBMT III	Excellent	Excellent	Sensitive to progression of memory impairment during PD course	• Brevity of administration • Very good psychometric properties	• Lacking normative datasets in some languages • Not available in all different languages	Recommended
Language						
BNT-60	Good	Excellent	NA	• Multiple normative data	• Not suitable to detect early PD-related language impairment • Cross-cultural differences • BNT-15 limited range for evaluating language impairment	Recommended
BNT-30				• Shorter but comparable to BNT-60		
BNT-15				• Very short		
Visuospatial function						
WAIS-III and WAIS-IV Matrix Reasoning	Excellent	Good	Good	• Very often used in clinical practice • Test is available in different languages • Robust normative datasets in different populations • Established factor structure including psychometric properties • No ceiling/floor effects of the subtest • Perceptual reasoning subtest is included in many PD research studies	• Validity of WAIS-IV subtest based on results of WAIS-III and WAIS-R • Validity of the perceptual reasoning index for PD is not known • Different versions of the battery	Recommended
Estimated premorbid intelligence						
NART/NAART/ NAART-R	High internal consistency	Excellent	NA	• Valid measure of premorbid intelligence in PD	• Illiterates or patients with reading disorders or significant visual impairment • Sensitive to different languages and cultures	Recommended

Abbreviations: RAVLT, Rey's Auditory Verbal Learning Test; PDD, Parkinson's disease dementia; CVLT, California Verbal Learning Test including the revised versions CVLT-II and CVLT3; PVLT, Philadelphia Verbal Learning Test; HVLT/HVLT-R, Hopkins Verbal Learning Test/Revised version; PD, Parkinson's disease; RBMT, Rivermead Behavioural Memory Test paragraph recall subtest including revised versions RBMT II, RBMT III; BNT-60, Boston Naming Test-60; BNT-30, Boston Naming Test-30; BNT-15, Boston Naming Test-15; NA, evidence not available; WAIS-III and WAIS-IV, Wechsler Adult Intelligence Scale III, IV; MR, Matrix Reasoning; NART National Adult Reading Test; NAART/NAART-R, North American Adult Reading Test-35/the North American Revision of the NART.

Test versions taken together as one test), four for memory, and one for visuospatial domains. Table 1 also includes premorbid intelligence measures (taken together as one test). Overall, based on the expert panel's full review of tests, 10 tests were "recommended with caveats," 11 "suggested," and 11 "listed"; their in-depth review and evaluation can be found in the Supporting Information.

Memory Domain

Rey's Auditory Verbal Learning Test (and Its Alternate Versions)

Scale description. The Rey's Auditory Verbal Learning Test (RAVLT) is a commonly used neuropsychological test of word list memory that uses 15 unrelated words. The original version, a one-trial word list, was developed by the Swiss psychologist, Édouard Claparède. It has been translated into English and copyrighted.[35] Further, it has been adapted into many other languages and modified by different groups, resulting in variability in procedures used across different studies.[36-38] More commonly, it consists of five learning trials with list A, each in a fixed order during learning, and then a free recall test, an interference trial with list B, subsequent recall of list A, delayed recall of list A after 20 or 30 minutes, and a recognition test.[39,40] There are normative data studies including regression-based-norms for different age groups and languages.[35,41-52] The primary results include learning, retention and recognition scores, and additional scores can be generated (e.g., process scores or error type).[53]

Strengths and weaknesses. RAVLT is a very well-established measure of memory functioning with a long history and psychometric analyses indicating good or excellent levels of reliability and validity (primarily delayed recall). The RAVLT is sensitive to memory impairment progression during PD course.

Recommendation level. The scale meets all required criteria and is recommended. The RAVLT is suitable for prevalence studies, for defining the level and profile of memory impairment, treatment trials, biomarker studies, and case–control studies.

California Verbal Learning Test, Including CVLT-II and CVLT3 and CVLT Short Form and Philadelphia Verbal Learning Test

Scale description. The California Verbal Learning Test (CVLT) is a multi-trial learning, recall and recognition word list test derived from the RAVLT.[38,40] The original version was created by Delis and colleagues[54] in 1987 with more recent refinements (CVLT-II, CVLT-short form [CVLT-SF], and CVLT3) and is copyrighted.[54-56] For the CVLT, the 16 items are explicitly drawn from four semantic categories (furniture, vegetables, animals, and transport/travelling), which differentiates it from the RAVLT. Performance on the whole test can yield a variety of memory indices.[54-56] In the late 1980s, Libon and colleagues[57] developed the CVLT-SF, published in 1996, which was a nine-word version with three categories (fruit, tools, and clothing) that was better tolerated by patients, especially those with severe neurological illness or dementia. The CVLT was also adapted as the Philadelphia Verbal Learning Test (PVLT) including nine- or 12-word lists.[58,59] The latest full version, CVLT3, includes the United States-based census data from 2015 and is stratified based on age, sex, race/ethnicity, education level, and geographic region.[60]

Strengths and weaknesses. CVLT is one of the most used and well-established measures of memory functioning with robust evidence of reliability and especially predictive validity for the presence of PD-MCI. It is sensitive to the progression of memory impairment in PD. The full CVLT may be difficult for patients with PD dementia. CVLT/CVLT-II/CVLT3/CVLT-SF are suitable for prevalence studies and also useful in clinical practice, treatment trials, and correlation with biological markers, but are not suitable for screening, as 12 words are too many.

Recommendation level. The scale meets all required criteria and is recommended. One of the most used and well-established measures of memory function with robust evidence of reliability and especially predictive validity for the development of PD-MCI in the long term. CVLT is sensitive to memory impairment progression during PD course.

Hopkins Verbal Learning Test, Including HVLT-Revised Version and Including Alternate Versions

Scale description. The original Hopkins Verbal Learning Test (HVLT) version was published by Brandt[61] in 1991. The test was modelled after the RAVLT, CVLT-SF, and PVLT and is copyrighted. The primary change from RAVLT and CVLT was the use of 12-word lists (four words drawn from each of three semantic categories, which are four-legged animals, precious stones, and human dwellings) plus the test contains six parallel forms, making it suitable for repeated assessment.[36] A delayed recall and recognition trial was introduced for the revised version (HVLT-R).[62] Demographically corrected normative data exist for several populations (eg, older, African-American, Latin American, Australian, Czech, Chinese and Portuguese).[62-69]

Strengths and weaknesses. HVLT can be used for repeated assessments with six parallel forms that are considered interchangeable.[70] HVLT is sensitive to the progression of cognitive impairment in PD; however, some reliability and validity indicators are only adequate. HVLT suffers from practice effects when normal individuals are given the same form after a 2-week interval.[71] HVLT/HVLT-R are suitable for prevalence studies and is also useful in clinical practice, treatment trials, and correlation with biological markers, but may not be suitable for screening because of the use of 12-word list.

Recommendation level. The scale meets all required criteria and is recommended. HVLT can be used for

repeated assessments as the six parallel forms are considered interchangeable. HVLT is sensitive to the progression of cognitive impairment in PD and has very good to excellent psychometric properties.

Rivermead Behavioural Memory Test, Story Recall Subtest (Including RBMT 2, RBMT 3)

Scale Description. The Rivermead Behavioural Memory Test (RBMT) battery was published in 1985 and was devised as an ecologically valid memory test and was updated in 2003 (RBMT 2) and 2008 (RBMT 3).[72-74] The subtest RBMT stories reflect a more common aspect of episodic memory compared to list learning.[14,36,75] The RBMT and RBMT-2 include 12 subtests and the RBMT-3 includes 14 subtests assessing aspects of visual and verbal recall and recognition, as well as immediate and delayed everyday memory. The story recall subtest consists of the auditory presentation of two news reports. Participants are asked to reproduce as many literal details as possible, both immediate (after each news report) and delayed. The scoring of the RBMT and RBMT-2 story recall includes two scoring systems: a screening score of 0 (fail) or 1 (pass) and a more detailed profile score of 0 (abnormal), 1 (borderline), or 2 (normal) both depending on the raw score for each news report (number of correct details), both immediate and delayed. The RBMT-3 scoring converts raw scores into a scaled score with a mean of 10 and a standard deviation of 3. Percentile ranks for scaled scores are also provided. All RBMT scores are normed with a healthy control group stratified for age and education[75,76] The test is also available in several languages (eg, Dutch, German, Arabic, Spanish and Chinese).

Strengths and weaknesses. RBMT stories (primarily delayed recall) are recommended by the IPMDS PD-MCI diagnostic criteria study group.[14] The scale has a short administration time and has excellent psychometric properties. RBMT has been used in many intervention studies, including the majority of deep brain stimulation studies.[77,78] However, the rights for adapting the scale may not be available from the supplier in each country. Additionally, determining change in individual cases is difficult because of incomplete reliable change indices.

Recommendation level. The scale meets all required criteria and is recommended. RBMT is suitable for screening, prevalence studies, differential reasoning (etiological), treatment trials, biomarker studies, and case–control studies.

Language Domain

Boston Naming Test-60; Boston Naming Test-30; Boston Naming Test-15

Scale description. Boston Naming Test (BNT)-15, BNT-30, and BNT-60 are tests of confrontation naming in which items shown as line drawings must be named spontaneously or after semantic or phonemic cueing. BNT-15, BNT-30, and BNT-60 are for clarity numbered separately and the BNT tests are considered not as an evolution of the test (such as CVLT/CVLT-2), but as different test versions. The most recent version is the second edition from 2001 (BNT-2).[79,80] BNT items were developed in 1983 and are identical for both versions, except that "noose" was replaced by "boomerang" in BNT-2.[81] Therefore, studies of the BNT are deemed equivalent to the BNT-2. There are also short-form versions of the original 60-item BNT, the BNT-30, and BNT-15.[82,83] The 30 and 15 refer to the number of stimulus items presented. There are three versions of the BNT-30 and four versions of the BNT-15 available. The original BNT and its versions are also copyrighted.

Strengths and weaknesses. BNT-15, BNT-30, and BNT-60 are the most used measure of confrontation naming. It is well-validated and used in virtually all patient populations in which cognitive assessment is indicated. Norms available for specific populations are available from several groups other than the test publisher.[36,80] Norms for older adult African Americans are available from Mayo's Older Americans Normative Studies.[84] PD studies suggest acceptable sensitivity to the change in heterogeneous cohorts. However, no difference between healthy controls and PD patients without dementia was detected, indicating that the scale may not be suitable for detecting language impairment in de novo patients or early stages of PD. On the other hand, patients with more severe cognitive impairment or dementia showed lower test performance compared with PD patients with no cognitive impairment.[85] Test performance seems to be unaffected by motor impairment. BNT-2 scores elicit ceiling effects and have a non-normal distribution of scores.[86] Because of the negative skew (also true for the BNT-30 and BNT-15), the BNT is most likely useful for identifying the presence of impairment rather than the level of impairment.[86-89] The BNT-15 version was recommended for use in PD.[90]

Recommendation level. BNT-2 is recommended because it meets all required criteria to assess naming abilities in PD patients. Sex-corrected normative values are recommended for score interpretation in all versions. Evidence for the use of BNT-2 and BNT-15 suggests that the BNT-30 version should also be suitable for PD.

Visuospatial Domain

Wechsler Adult Intelligence Scale: Matrix Reasoning

Scale description. This subtest of the Wechsler Adult Intelligence Scale (WAIS) perceptual reasoning domain evaluates nonverbal abstract problem-solving and inductive reasoning behind visuospatial elements and is also considered a measure of fluid intelligence.[91] The test contains 26 items differentiated into four types of nonverbal

reasoning tasks: pattern completion, classification, analogy, and serial reasoning. The examinee views an array of pictures with one missing square and selects the picture that fits the array from five options (maximum = 26 points). Test performance is proposed to be culture and language-free, with no time limit.[92]

Strengths and weaknesses. The Matrix Reasoning (MR) is a core subtest of the perceptual reasoning index scale, with MR included in the WAIS-III and WAIS-IV versions. It shows good reliability. In PD, the criterion validity of the WAIS-IV MR is not well explored, but has been evaluated for previous test versions (WAIS-III, WAIS-R). There is a high correspondence between WAIS-III and WAIS-IV.[93,94] The MR can be applied in early and late PD disease stages. The sensitivity of change for MR[95] is good. However, further investigation for the WAIS-IV tests perceptual reasoning subtests is needed.

Recommendation level. The MR WAIS-IV subtest is recommended to assess perceptual reasoning in PD. The WAIS-IV plays a central role in clinical practice. It has broad applicability to individuals ages 16 to 89 years. The reliability of the perceptual reasoning subscale is good. The internal structure of the subtest has been confirmed by factor analysis.

Premorbid Intelligence Domain

National Adult Reading Test; North American Adult Reading Test-35

Scale description. The 50-item National Adult Reading Test (NART), 61-item North American Adult Reading Test (NAART)-Revised (NAART-R) (the North American Revision of the NART), and the NART-R United Kingdom (UK revision) have been used most frequently. There are also three abbreviated forms, the 17-Item NART (NART 17), the Mini-NART (23 items) and the Short NART (which is based on the first half of the NART) as well as the Cambridge Contextual Reading Test that uses NART words embedded in sentences to provide context for the examinee.[96-99] The NART's 50 phonetically irregular words have graded levels of difficulty, and accurate reading of these words is used to estimate premorbid intellectual ability. The total score is the sum of all items that are pronounced incorrectly. The fewer the number of incorrectly pronounced words, the higher the estimate of premorbid intellectual ability, although estimates are poorer for intelligent quotients (IQs) in the more extreme ranges.[100] The test is copyrighted by H. Nelson and is in the public domain. The scale has several language adaptations into Dutch, French, Japanese, Swedish, and Czech, and the original NART was adapted for Australian English or American English versions (AMNART) and may need modification for some non-UK English–speaking countries.[101-109]

Strengths and weaknesses. The NART/NAART are quick measures to estimate premorbid verbal intelligence free of sensitivity to early dementia stages.[34] The use of the NART to estimate premorbid cognition changes the proportion of patients diagnosed with level 1 IPMDS PD-MCI criteria compared with assessment without this information to minimize the influence of premorbid cognitive ability and cognitive reserve concerning current cognitive status.[27] The addition of a premorbid IQ measure may complement other metrics used to diagnose PD-MCI.[14]

Recommendation level. The NART/NAART tests meet the required criteria to assess premorbid cognitive ability in PD patients and is recommended. These measures are intended to be resistant to moderate levels of neurodegeneration so criteria regarding change are not applicable in this instance. Overall, NART/NAART is suitable for the assessment of premorbid verbal intellectual ability.

Discussion and Recommendations

This review provides critique and recommendations of tests that assess memory, language, and visuospatial domains, as well as premorbid intelligence, across all cognitive stages in PD (PD with normal cognition, PD-MCI, and PDD).[14,15,110] The recommended tests show robust psychometric evidence that makes them suitable for differentiating PD-MCI and PDD from patients with normal cognition. Tests recommended with caveats, suggested, or listed are provided in Supporting Information. The review summarizes the key psychometric properties of those tests listed as recommended, including their validity, reliability, classification accuracy, and sensitivity to change, as well as their clinimetric limitations. Moreover, it provides information about recent developments in the applicability of these instruments for neuropsychological assessment in PD.

For consistency, we adopted the classification into domains introduced by Litvan et al.[14] and grouped these tests into Memory, Language, and Visuospatial function domains (with attention/working memory and executive functions being assessed in a parallel review).

Memory

In general, word list memory tests, such as RAVLT/CVLT/HVLT-R that tap key memory processes (encoding/retention/recognition) are sensitive with appropriate psychometric indicators of validity, reliability, and sensitivity to change in PD. They vary in the use of free recall only (RAVLT), free recall and clustering (HVLT-R), free recall and clustering plus cued recall as in CVLT-3 and related versions. Similar findings apply to prose recall measures with a delayed recall condition, such as RBMT-III or earlier versions of the story recall subtest. Further clinimetric analysis should be directed at promising tests minimizing the role of the executive and attentional dysfunction in PD by using controlled learning

and recall paradigms, such as Selective Reminding Test (SRT) or Memory Binding Test (MBT). Controlled learning assures equal attention processing of all items, shows that individuals can identify items by their cues, induces all individuals to do the same processing, and ensures that low recall is because of impairment of memory (not limited by attention, strategies, or depression) whereas controlled cued recall assures attention to and equal testing of all items, controls order of recall, all subjects recall all items in the same order, equalizes the interval between learning and recall, obviates need for interference before recall, and prevents output order effects.

In the visuospatial memory domain, the Rey-Osterrieth Complex Figure Test (ROCFT) is "suggested." However, of note is that the ROCFT cannot minimize graphomotor or attentional-executive impairments and their negative impact on visuospatial memory performance (see also Table S1 for tests recommended with caveats, such as WMS-IV logical memory or BVMT-R).

Future directions in research of memory tests in PD should be directed to demonstrate a correlation with PD cognitive biomarkers and their role in the detection of cognitive deficits in the preclinical stages of PD and other synucleinopathies. The advantages and disadvantages of computer-based versions of the instruments and computational capacities of digital technologies are not fully understood and need to be examined in PD specifically.[111]

Language

In this domain, the BNT (including 60-, 15-, or 30-item versions) is recommended, especially for more severely impaired patients. However, additional research needs to determine if language measures can be found that are sensitive to preclinical or early stages of PD.[112,113] Verbal fluency tests can be considered tests of either language or executive function. For our purposes, they are covered as tests of executive function in the companion review.

Overall, the major problem in language tests recommended for use in PD will be to limit motor confounds (i.e., hypokinetic dysarthria) or relative insensitivity of these measures to de novo or early stages of PD. In future, a very dynamic evolution of scales for the measurement of speech and language difficulties in PD with the use of artificial intelligence (AI) and computational processing of words and phrases can be expected including their correlation with biomarkers in preclinical stages of PD.[114]

Visuospatial Function

This broad domain consists of visuo-spatial, visuo-perceptive, and visuo-constructive tests. We recommend the WAIS-IV MR subtest (or earlier versions). See also Table S1 for memory and visuospatial tests recommended with caveats, such as WAIS-IV Block Design, ROCFT or Benton Judgment of Line Orientation (B-JoLO).

In sum, these measures are confronted in PD more (ROCFT) or less (MR, B-JoLO) with the constraint of how to limit motor or graphomotor confounds.[111] Prospectively, their correlation with biomarkers and improvements with cognitive enhancing treatment must be established. The computational power of digital technologies in these tests should be developed for PD patients.[111]

Premorbid Intelligence Estimation

These measures provide an estimate of premorbid IQ only and not for language per se. However, the premorbid intelligence measures play a significant role by delineating the cognitive potential of patients and were recommended.

It remains an open question whether other premorbid intelligence measures, such as the Wechsler Test of Adult Reading (WTAR) or the Test of Premorbid Functioning (TOPF) will show better discriminative potential than the NART or if these measures will be replaced by digitally assisted technology using premorbid data long before disease onset with perhaps more robust estimates and higher ecological validity.

The limits of the current review must be fully acknowledged. PD progression and disease stage are two of the key principles regarding test selection for neuropsychological assessment of PD.[26,115,116] None of the "Recommended" tests are equally sensitive to PD progression from the stage with no detectable cognitive impairment to PDD. "Recommended" tests may have differential sensitivity to PD-MCI phenotypes, including risky cognitive profiles associated with progression to PDD.[117] Moreover, the current review was not dedicated to the prodromal stage of PD as represented by isolated rapid eye movement (REM) sleep behavior disorder (iRBD),[118] which has in recent years become one of the prognostic factors regarding premorbid cognitive decline in α-synuclein disorders and may provide novel directions for test development.[119,120] In addition, every test taps different processes (e.g., does the "memory" impairment in PD involve primarily poor encoding, poor retention, and/or poor retrieval?; does the "language" deficits in PD primarily involve word production, retrieval, comprehension or hypokinetic dysarthria?). Indeed, in "Recommended" memory tests "acquisition" could be as relevant as "retention" and "retrieval" or "susceptibility to interference effects" in their ability to distinguish between memory deficits. These questions are not captured by our review. More specifically, no tests of visuospatial abilities had high enough validity or reliability, or had not been used with PD cohorts, to reach the "Recommended" level. It can be argued that WAIS MR as the recommended measure of visuospatial ability does not capture core visuospatial function because it also involves abstract and inductive reasoning skills (i.e., aspects of executive function). New normative data using current tests and the

development of new tests in the visuospatial domain should be considered and the tests should be validated for use in PD.

Overall, all tests reviewed suffer from legal constraints, which means that many tests are not available in many countries because of unresolved legal or business issues, which result in the absence of translations from English. Another important point is that copyright is one issue, and cost is another. From seven (not counting NART) tests recommended, the cost associated with the use of each starts from the lowest BNT-2 for $145.00 up to $1599 necessary for the complete kit of WAIS-IV, which may restrict their use in clinical care and research.

One should also consider that the PD-MCI and PDD criteria defined by the IPMDS differ in some respects from other neuropsychiatric definitions (e.g., minor and major neurocognitive disorders as defined by Diagnostic and Statistical Manual of Mental Disorders, fifth edition).

Our review sought to identify neuropsychological tests most suitable for the assessment of PD cognitive impairment at different stages (PD with normal cognition/PD-MCI/PDD)[13,14], thereby extending the previous IPMDS review on global scales for cognitive screening.[32] Specific neuropsychological tests often provide a better balance between sensitivity, specificity, and diagnostic accuracy and are generally preferred over the global cognitive screening scales, such as the Montreal Cognitive Assessment or Mini-Mental State Examination.[13] These neuropsychological tests have a high predictive value in the detection of PD-MCI or PDD, minimizing false positive rates. They enable the clinician to convert the raw scores to interpretable normed scores according to age, education, sex, and race given normative data that are available for these instruments in different populations and countries.

In conclusion, after undertaking an in-depth review of the advantages and disadvantages of cognitive tests, including their psychometric properties in memory, language, and visuospatial domains, we present a list of the recommended tests for the assessment of PD cognition across stages of impairment. In general, their sensitivity to cognitive enhancing treatment, their correlation with PD cognitive biomarkers, and their transformation into computer-based versions using digital technology all await improvement. The current review underlines a strong need for further evidence on existing and new instruments for the emerging era of translational neuroscience for PD patients. ■

Author Roles: O.B.: conception, design, execution, analysis, and writing and revising the manuscript. S.B.: design, execution, analysis, and writing of the manuscript. D.M.C.: design, execution, analysis, and writing of the manuscript. B.C.: design, execution, analysis, and writing of the manuscript. A.C.-G.: design, execution, analysis, and writing of the manuscript. J.C.D.-A.: design, execution, analysis, and writing of the manuscript. A.D.: design, execution, analysis, and writing of the manuscript. R. F.: design, execution, analysis, and writing of the manuscript. A.G.: design, execution, analysis, and writing of the manuscript. H.E.: design, execution, analysis, and writing of the manuscript. H.H.: design, execution, analysis, and writing of the manuscript. J.K.: design, execution, analysis, and writing of the manuscript. B.L.: design, execution, analysis, and writing of the manuscript. I.L.-S.: design, execution, analysis, and writing of the manuscript. M.S.: design, execution, analysis, and writing of the manuscript. T.A.M.: execution and editing of the manuscript. Á.S.-F.: execution and editing of the manuscript. M.H.S.T.: execution and editing of the manuscript. M.S.: design, execution, analysis, and writing of the manuscript. D.W.: conception, design, execution, analysis, and writing and revising the manuscript. R.B.: conception, design, execution, analysis, and writing and revising the manuscript. G.J.G.: conception, design, execution, analysis, and writing and revising the manuscript.

Acknowledgments: We would express many thanks to the following colleagues. Truley Juneau, M.A., Holly Klecha, M.A., and Bridget Logan, M.A. for help during the preparation of the manuscript and reviews, also to the MDS COA program secretariat. Open access publishing facilitated by Univerzita Karlova, as part of the Wiley - CzechELib agreement.

Disclosure

Financial disclosure all authors: The authors declare that there are no additional disclosures to report.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. Poewe W, Seppi K, Tanner CM, et al. Parkinson disease. Nat Rev Dis Prim 2017;3(1):17013.
2. Schrag A, Jahanshahi M, Quinn N. What contributes to quality of life in patients with Parkinson's disease? J Neurol Neurosurg Psychiatry 2000;69(3):308–312.
3. Aarsland D, Larsen JP, Karlsen K, Lim NG, Tandberg E. Mental symptoms in Parkinson's disease are important contributors to caregiver distress. Int J Geriatr Psychiatry 1999;14(10):866–874.
4. Aarsland D, Brønnick K, Ehrt U, et al. Neuropsychiatric symptoms in patients with Parkinson's disease and dementia: frequency, profile and associated care giver stress. J Neurol Neurosurg Psychiatry 2007;78(1):36–42.
5. Pirogovsky E, Schiehser DM, Obtera KM, et al. Instrumental activities of daily living are impaired in Parkinson's disease patients with mild cognitive impairment. Neuropsychology 2014;28(2):229–237.
6. Becker S, Bode M, Brockmann K, et al. Cognitive-driven activities of daily living impairment as a predictor for dementia in Parkinson disease: a longitudinal cohort study. Neurology 2022;99(23):e2548–e2560.
7. Aarsland D, Batzu L, Halliday GM, et al. Parkinson disease-associated cognitive impairment. Nat Rev Dis Primers 2021;7(1):47.
8. Postuma RB, Iranzo A, Hu M, et al. Risk and predictors of dementia and parkinsonism in idiopathic REM sleep behaviour disorder: a multicentre study. Brain 2019;142(3):744–759.
9. Williams-Gray CH, Mason SL, Evans JR, et al. The CamPaIGN study of Parkinson's disease: 10-year outlook in an incident population-based cohort. J Neurol Neurosurg Psychiatry 2013;84(11):1258–1264.
10. Broeders M, de Bie RM, Velseboer DC, Speelman JD, Muslimovic D, Schmand B. Evolution of mild cognitive impairment in Parkinson disease. Neurology 2013;81(4):346–352.
11. Muslimovic D, Post B, Speelman JD, Schmand B. Cognitive profile of patients with newly diagnosed Parkinson disease. Neurology 2005;65(8):1239–1245.

12. Berg D, Postuma RB, Adler CH, et al. MDS research criteria for prodromal Parkinson's disease. Mov Disorders 2015;30(12): 1600–1611.
13. Boel JA, de Bie RMA, Schmand BA, et al. Level I PD-MCI using global cognitive tests and the risk for Parkinson's disease dementia. Mov Disord Clin Pract 2022;9(4):479–483.
14. Litvan I, Goldman JG, Troster AI, et al. Diagnostic criteria for mild cognitive impairment in Parkinson's disease: Movement Disorder Society task force guidelines. Mov Disord 2012;27(3):349–356.
15. Emre M, Aarsland D, Brown R, et al. Clinical diagnostic criteria for dementia associated with Parkinson's disease. Mov Disord 2007;22(12):1689–1707. quiz 1837
16. Hely MA, Reid WG, Adena MA, Halliday GM, Morris JG. The Sydney multicenter study of Parkinson's disease: the inevitability of dementia at 20 years. Mov Disord 2008;23(6):837–844.
17. Yarnall AJ, Breen DP, Duncan GW, et al. Characterizing mild cognitive impairment in incident Parkinson disease: the ICICLE-PD study. Neurology 2014;82(4):308–316.
18. Aarsland D, Brønnick K, Fladby T. Mild cognitive impairment in Parkinson's disease. Curr Neurol Neurosci Rep 2011;11(4): 371–378.
19. Saredakis D, Collins-Praino LE, Gutteridge DS, Stephan BCM, Keage HAD. Conversion to MCI and dementia in Parkinson's disease: a systematic review and meta-analysis. Parkinsonism Relat Disord 2019;65:20–31.
20. Gratwicke J, Jahanshahi M, Foltynie T. Parkinson's disease dementia: a neural networks perspective. Brain 2015;138(Pt 6):1454–1476.
21. Bezdicek O, Ballarini T, Buschke H, et al. Memory impairment in Parkinson's disease: the retrieval versus associative deficit hypothesis revisited and reconciled. Neuropsychology 2019;33(3):391–405.
22. Bezdicek O, Ballarini T, Růžička F, et al. Mild cognitive impairment disrupts attention network connectivity in Parkinson's disease: a combined multimodal MRI and meta-analytical study. Neuropsychologia 2018;112:105–115.
23. Kehagia AA, Barker RA, Robbins TW. Neuropsychological and clinical heterogeneity of cognitive impairment and dementia in patients with Parkinson's disease. Lancet Neurol 2010;9(12):1200–1213.
24. Pagonabarraga J, Kulisevsky J. Cognitive impairment and dementia in Parkinson's disease. Neurobiol Dis 2012;46(3):590–596.
25. Czernecki V, Benchetrit E, Houot M, et al. Social cognitive impairment in early Parkinson's disease: a novel "mild impairment"? Parkinsonism Relat Disord 2021;85:117–121.
26. Hoogland J, Boel JA, de Bie RMA, et al. Mild cognitive impairment as a risk factor for Parkinson's disease dementia. Mov Disord 2017;32(7):1056–1065.
27. Marras C, Armstrong MJ, Meaney CA, et al. Measuring mild cognitive impairment in patients with Parkinson's disease. Mov Disord 2013;28(5):626–633.
28. Goldman JG, Holden S, Bernard B, Ouyang B, Goetz CG, Stebbins GT. Defining optimal cutoff scores for cognitive impairment using Movement Disorder Society task force criteria for mild cognitive impairment in Parkinson's disease. Mov Disord 2013; 28(14):1972–1979.
29. Meissner WG, Frasier M, Gasser T, et al. Priorities in Parkinson's disease research. Nat Rev Drug Discov 2011;10(5):377–393.
30. Stocchi F, Olanow CW. Neuroprotection in Parkinson's disease: clinical trials. Ann Neurol 2003;53(S3):S87–S99.
31. Stebbins GT. Cognitive impairment screening scales. In: Sampaio C, Goetz CG, Schrag A, eds. Rating Scales in Parkinson's Disease. Oxford: Oxford University Press; 2012.
32. Skorvanek M, Goldman JG, Jahanshahi M, et al. Global scales for cognitive screening in Parkinson's disease: critique and recommendations. Mov Disord 2018;33(2):208–218.
33. Schrag A, Barone P, Brown RG, et al. Depression rating scales in Parkinson's disease: critique and recommendations. Mov Disord 2007;22(8):1077–1092.
34. Szeto JY, Mowszowski L, Gilat M, Walton CC, Naismith SL, Lewis SJ. Assessing the utility of the Movement Disorder Society task force level 1 diagnostic criteria for mild cognitive impairment in Parkinson's disease. Parkinsonism Relat Disord 2015;21(1): 31–35.
35. Schmidt M. Rey Auditory and Verbal Learning Test: A Handbook. Los Angeles, CA: Western Psychological Services; 1996.
36. Sherman E, Tan J, Hrabok M. A Compendium of Neuropsychological Tests. Fundamentals of Neuropsychological Assessment and Test Reviews for Clinical Practice. 4th ed. New York: Oxford University Press; 2023.
37. Mitrushina M, Boone KB, Razani J, D'Elia LF. Handbook of Normative Data for Neuropsychological Assessment. New York: Oxford University Press; 2005.
38. Rey A. L'examen Clinique en Psychologie. Paris: Presses Universitaries de France; 1964.
39. Boake C. Édouard Claparède and the auditory verbal learning test. J Clin Exp Neuropsyc 2000;22(2):286–292.
40. Rey A. L'examen psychologique dans les cas d'encephopathie traumatique (the psychological examination of cases of traumatic encephalopathy). Arch Psychol 1941;28:286–340.
41. Ivnik RJ, Malec JF, Tangalos EG, Petersen RC, Kokmen E, Kurland LT. The auditory-verbal learning test (AVLT): norms for ages 55 years and older. Psychol Assess: J Consulting Clin Psychol 1990;2(3):304–312.
42. Ivnik RJ, Malec JF, Smith GE, et al. Mayo's older Americans normative studies: updated AVLT norms for ages 56 to 97. Clin Neuropsychol 1992;6(Suppl):83–104.
43. Bezdicek O, Stepankova H, Motak L, et al. Czech version of Rey auditory verbal learning test: normative data. Aging Neuropsychol Cognit 2014;21(6):693–721.
44. Stricker NH, Christianson TJ, Lundt ES, et al. Mayo normative studies: regression-based normative data for the auditory verbal learning test for ages 30-91 years and the importance of adjusting for sex. J Int Neuropsychol Soc 2021;27(3):211–226.
45. Ferreira Correia A, Campagna OI. The Rey auditory verbal learning test: normative data developed for the Venezuelan population. Arch Clin Neuropsychol 2013;29(2):206–215.
46. Vakil E, Greenstein Y, Blachstein H. Normative data for composite scores for children and adults derived from the Rey auditory verbal learning test. Clin Neuropsychol 2010;24(4):662–677.
47. Messinis L, Nasios G, Mougias A, et al. Age and education adjusted normative data and discriminative validity for Rey's auditory verbal learning test in the elderly Greek population. J Clin Exp Neuropsychol 2016;38(1):23–39.
48. Lee T, Yuen K, Chan C. Normative data for neuropsychological measures of fluency, attention, and memory measures for Hong Kong Chinese. J Clin Exp Neuropsychol 2002;24(5):615–632.
49. Helmstaedter C, Durwen HF. VLMT: Verbaler Lern- und Merkfähigkeitstest: Ein praktikables und differenziertes Instrumentarium zur Prüfung der verbalen Gedächtnisleistungen. [VLMT: A useful tool to assess and differentiate verbal memory performance.]. Schweizer Arch Neurol Neurochir Psychiatr 1990;141(1):21–30.
50. Van der Elst W, van Boxtel MP, van Breukelen GJ, Jolles J. Rey's verbal learning test: normative data for 1855 healthy participants aged 24–81 years and the influence of age, sex, education, and mode of presentation. J Int Neuropsychol Soc 2005;11(3):290–302.
51. Caltagirone C, Gainotti G, Carlesimo GA, Parnetti L. Batteria per la valutazione del deterioramento mentale: I. Descrizione di uno strumento di diagnosi neuropsicologica. [The Mental Deterioration Battery: I. Description of a neuropsychological diagnostic instrument.]. Arch Psicol Neurol Psichiatr 1995;56(4):461–470.
52. Rezvanfard M, Ekhtiari H, Noroozian M, Rezvanifar A, Nilipour R, Karimi JG. The Rey auditory verbal learning test: alternate forms equivalency and reliability for the Iranian adult population (Persian version). Arch Iran Med 2011;14(2):104–109.
53. Woodard JL, Dunlosky JA, Salthouse TA. Task decomposition analysis of intertrial free recall performance on the Rey auditory verbal learning test in normal aging and Alzheimer's disease. J Clin Exp Neuropsychol 1999;21(5):666–676.

54. Delis DC, Kramer JH, Kaplan E, Ober BA. The California Verbal Learning Test: Research Edition, Adult Version. San Antonio, TX: The Psychological Corporation; 1987.
55. Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test – Second Edition, Adult Version. San Antonio, TX: The Psychological Corporation; 2000.
56. Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test – Third Edition, Adult Version. San Antonio, TX: The Psychological Corporation; 2017.
57. Libon DJ, Mattson RE, Glosser G, et al. A nine-word dementia version of the California verbal learning test. Clin Neuropsychol 1996;10(3):237–244.
58. Price CC, Garrett KD, Jefferson AL, et al. Leukoaraiosis severity and list-learning in dementia. Clin Neuropsychol 2009;23(6): 944–961.
59. Bezdicek O, Libon DJ, Stepankova H, et al. Development, validity, and normative data study for the 12-word Philadelphia verbal learning test [czP(r)VLT-12] among older and very old Czech adults. Clin Neuropsychol 2014;28(7):1162–1181.
60. Farrer TJ, Drozdick LW. Essentials of the California Verbal Learning Test: CVLT-C, CVLT-2, & CVLT3. Hoboken, NJ: John Wiley & Sons; 2020.
61. Brandt J. The Hopkins verbal learning test: development of a new memory test with six equivalent forms. Clin Neuropsychol 1991; 5(2):125–142.
62. Brandt J, Benedict RHB. Hopkins Verbal Learning Test–Revised. Odessa, Fla: Psychological Assessment Resources; 2001.
63. Vanderploeg RD, Schinka JA, Jones T, Small BJ, Graves AB, Mortimer JA. Elderly norms for the Hopkins verbal learning test-revised. Clin Neuropsychol 2000;14(3):318–324.
64. Duff K. Demographically corrected normative data for the Hopkins verbal learning test-revised and brief visuospatial memory test-revised in an elderly sample. Appl Neuropsychol Adult 2016;23(3): 179–185.
65. Hester RL, Kinsella GJ, Ong B, Turner M. Hopkins verbal learning test: normative data for older Australian adults. Aust Psychol 2004;39(3):251–255.
66. Arango-Lasprilla JC, Rivera D, Garza MT, et al. Hopkins verbal learning test- revised: normative data for the Latin American Spanish speaking adult population. NeuroRehabilitation 2015;37(4):699–718.
67. Havlik F, Michalec J, Kališová L, et al. The normative data study of the Czech MATRICS consensus cognitive battery. Clin Neuropsychol 2021;35(sup1):S50–s64.
68. Shi J, Tian J, Wei M, Miao Y, Wang Y. The utility of the Hopkins verbal learning test (Chinese version) for screening dementia and mild cognitive impairment in a Chinese population. BMC Neurol 2012;12:136.
69. Vicente SG, Ramos-Usuga D, Barbosa F, et al. Regression-based norms for the Hopkins verbal learning test-revised and the Rey-Osterrieth complex figure in a Portuguese adult population. Arch Clin Neuropsychol 2021;36(4):587–596.
70. Benedict RHB, Schretlen D, Groninger L, Brandt J. Hopkins verbal learning test—revised: normative data and analysis of inter-form and test–retest reliability. Clin Neuropsychol 1998;12(1):43–55.
71. Benedict RH, Zgaljardic DJ. Practice effects during repeated administrations of memory tests with and without alternate forms. J Clin Exp Neuropsychol 1998;20(3):339–352.
72. Wilson B, Cockburn J, Baddeley AD. Rivermead Behavioural Memory Test. Flempton: Thames Valley Test Company; 1985.
73. Wilson B, Cockburn J, Baddeley A. The Rivermead Behavioural Memory Test. 2nd ed. London: Pearson Assessment; 2003.
74. Wilson BA. The Rivermead Behavioural Memory Test - Third Edition RBMT 3. 3rd ed. London: Pearson Assessment; 2008.
75. Agelink van Rentergem JA, de Vent NR, Schmand BA, Murre JMJ, Staaks JPC, Huizenga HM. The factor structure of cognitive functioning in cognitively healthy participants: a meta-analysis and meta-analysis of individual participant data. Neuropsychol Rev 2020;30(1):51–96.
76. de Vent NR, Agelink van Rentergem JA, Schmand BA, Murre JM, Huizenga HM. Advanced neuropsychological diagnostics infrastructure (ANDI): a normative database created from control datasets. Front Psychol 2016;7:1601.
77. Odekerken VJ, van Laar T, Staal MJ, et al. Subthalamic nucleus versus globus pallidus bilateral deep brain stimulation for advanced Parkinson's disease (NSTAPS study): a randomised controlled trial. Lancet Neurol 2013;12(1):37–44.
78. Boel JA, Odekerken VJ, Schmand BA, et al. Cognitive and psychiatric outcome 3 years after globus pallidus pars interna or subthalamic nucleus deep brain stimulation for Parkinson's disease. Parkinsonism Relat Disord 2016;33:90–95.
79. Kaplan E, Goodglass H, Barresi B. Boston Naming Test–Second Edition. 2nd ed. Austin, TX: Pro-Ed; 2001.
80. Strauss E, Sherman EMS, Spreen O. A Compendium of Neuropsychological Tests: Administration, Norms, and Commentary. 3rd ed. New York, NY, US: Oxford University Press; 2006.
81. Zimmerman D, Attridge J, Rolin S, Davis J. Psychometric equivalence of standard and prorated Boston naming test scores. Assessment 2020;29(3):527–534.
82. Williams BW, Mack W, Henderson VW. Boston naming test in Alzheimer's disease. Neuropsychologia 1989;27(8):1073–1079.
83. Mack WJ, Freed DM, Williams BW, Henderson VW. Boston naming test: shortened versions for use in Alzheimer's disease. J Gerontol 1992;47(3):P154–P158.
84. Lucas JA, Ivnik RJ, Smith GE, et al. Mayo's older African Americans normative studies: norms for Boston naming test, controlled Oral word association, category fluency, animal naming, token test, WRAT-3 Reading, trail making test, Stroop test, and judgment of line orientation. Clin Neuropsychol 2005;19(2):243–269.
85. Tröster AI, Stalp LD, Paolo AM, Fields JA, Koller WC. Neuropsychological impairment in Parkinson's disease with and without depression. Arch Neurol 1995;52(12):1164–1169.
86. Harry A, Crowe SF. Is the Boston naming test still fit for purpose? Clin Neuropsychol 2014;28(3):486–504.
87. Mitrushina M, Boone KB, Razani J, D'Elia LF. Handbook of Normative Data for Neuropsychological Assessment. 2nd ed. Oxford: Oxford University Press; 2005.
88. Bezdicek O, Rosická AM, Mana J, Libon DJ, Kopeček M, Georgi H. The 30-item and 15-item Boston naming test Czech version: item response analysis and normative values for healthy older adults. J Clin Exp Neuropsychol 2021;43(9):890–905.
89. Pedraza O, Sachs BC, Ferman TJ, Rush BK, Lucas JA. Difficulty and discrimination parameters of Boston naming test items in a consecutive clinical series. Arch Clin Neuropsychol 2011;26(5): 434–444.
90. Watson GS, Cholerton BA, Gross RG, et al. Neuropsychologic assessment in collaborative Parkinson's disease research: a proposal from the National Institute of Neurological Disorders and Stroke Morris K. Udall centers of excellence for Parkinson's disease research at the University of Pennsylvania and the University of Washington. Alzheimers Dement 2013;9(5):609–614.
91. Ryan JJ, Lopez SJ. Wechsler Adult Intelligence Scale-III. In: Dorfman WI, Hersen M, eds. Understanding Psychological Assessment. New York, NY: Kluwer Academic/Plenum Publishers; 2001.
92. von Aster M, Neubauer A, Horn R, eds. Wechsler Intelligenztest für Erwachsene (WIE); Deutschsprachige Bearbeitung und Adaptation des WAIS-III von David Wechsler. Frankfurt am Main: Pearson Assessment; 2006.
93. Petermann F, ed. Wechsler Adult Intelligence Scale (WAIS-IV). 4th ed. Frankfurt am Main: Deutschsprachige Adaptation der WAIS-IV von D. Wechsler: Grundlagen, Testauswertung und Interpretation. Pearson Assessment; 2012.
94. Valentine T, Block C, Eversole K, Boxley L, Dawson E. Wechsler adult intelligence scale-IV (WAIS-IV). Clinical indications and consideration. In: Carducci BJ, Nave CS, eds. The Wiley Encyclopedia of Personality and Individual Differences. Hoboken, NJ: John Wiley & Sons; 2020:457–464.
95. Odekerken VJ, Boel JA, Geurtsen GJ, et al. Neuropsychological outcome after deep brain stimulation for Parkinson disease. Neurology 2015;84(13):1355–1361.

96. Nelson HE. National Adult Reading Test (NART): For the Assessment of Premorbid Intelligence in Patients with Dementia: Test Manual. Windsor: NFER-Nelson; 1982.
97. Nelson HE, Willison JR. The Revised National Adult Reading Test: Test Manual. Windsor: NFER-Nelson, Windsor; 1991.
98. Blair JR, Spreen O. Predicting premorbid IQ: a revision of the national adult reading test. Clin Neuropsychol 1989;3(2):129–136.
99. Uttl B. North American adult Reading test: age norms, reliability, and validity. J Clin Exp Neuropsychol 2002;24(8):1123–1137.
100. Bright P, Hale E, Gooch VJ, Myhill T, van der Linde I. The National Adult Reading Test: restandardisation against the Wechsler adult intelligence scale-fourth edition. Neuropsychol Rehabil 2018;28(6):1019–1027.
101. Schmand B, Bakker D, Saan R, Louman J. The Dutch Reading test for adults: a measure of premorbid intelligence level. Tijdschr Gerontol Geriatr 1991;22(1):15–19.
102. Mackinnon A, Ritchie K, Mulligan R. The measurement properties of a French language adaptation of the National Adult Reading Test. Int J Methods Psychiatr Res 1999;8(1):27–38.
103. Hirata-Mogi S, Koike S, Toriyama R, Matsuoka K, Kim Y, Kasai K. Reliability of a paper-and-pencil version of the Japanese adult Reading test short version. Psychiatry Clin Neurosci 2016; 70(8):362.
104. Rolstad S, Nordlund A, Gustavsson MH, et al. The Swedish National Adult Reading Test (NART-SWE): a test of premorbid IQ. Scand J Psychol 2008;49(6):577–582.
105. Halliday TJ. The Development of a New Zealand Adult Reading Test [Masters]. Hamilton, New Zealand: University of Waikato The University of Waikato; 2006.
106. Watt S, Ong B, Crowe SF. Developing a regression equation for predicting premorbid functioning in an Australian sample using the National Adult Reading Test. Aust J Psychol 2018;70(2):186–195.
107. Krámská L. Assessment of Premorbid Intellect in Neuropsychology. Czech Reading Test. Otrokovice: Propsyco; 2014.
108. Mathias JL, Bowden SC, Barrett-Woodbridge M. Accuracy of the Wechsler test of adult Reading (WTAR) and National Adult Reading Test (NART) when estimating IQ in a healthy Australian sample. Aust Psychol 2007;42(1):49–56.
109. Grober E, Sliwinski M. Development and validation of a model for estimating premorbid verbal intelligence in the elderly. J Clin Exp Neuropsychol 1991;13(6):933–949.
110. Dubois B, Burn D, Goetz C, et al. Diagnostic procedures for Parkinson's disease dementia: recommendations from the movement disorder society task force. Mov Disorders 2007;22(16): 2314–2324.
111. Burn D, Weintraub D, Ravina B, Litvan I. Cognition in movement disorders: where can we hope to be in ten years? Mov Disord 2014;29(5):704–711.
112. Šubert M, Novotný M, Tykalová T, et al. Spoken language alterations can predict Phenoconversion in isolated rapid eye movement sleep behavior disorder: a multicenter study. Ann Neurol 2024; 95(3):530–543.
113. Rusz J, Cmejla R, Tykalova T, et al. Imprecise vowel articulation as a potential early marker of Parkinson's disease: effect of speaking task. J Acoust Soc Am 2013;134(3):2171–2181.
114. Miglis MG, Adler CH, Antelmi E, et al. Biomarkers of conversion to α-synucleinopathy in isolated rapid-eye-movement sleep behaviour disorder. Lancet Neurol 2021;20(8):671–684.
115. Biundo R, Weis L, Antonini A. Cognitive decline in Parkinson's disease: the complex picture. NPJ Parkinsons Dis 2016;2:16018.
116. Wallace ER, Segerstrom SC, van Horne CG, Schmitt FA, Koehl LM. Meta-analysis of cognition in Parkinson's disease mild cognitive impairment and dementia progression. Neuropsychol Rev 2022;32(1):149–160.
117. Biundo R, Weis L, Facchini S, et al. Cognitive profiling of Parkinson disease patients with mild cognitive impairment and dementia. Parkinsonism Relat Disord 2014;20(4):394–399.
118. Postuma RB, Bertrand JA, Montplaisir J, et al. Rapid eye movement sleep behavior disorder and risk of dementia in Parkinson's disease: a prospective study. Mov Disorders 2012;27(6):720–726.
119. Mao J, Huang X, Yu J, et al. Association between REM sleep behavior disorder and cognitive dysfunctions in Parkinson's disease: a systematic review and meta-analysis of observational studies. Front Neurol 2020;11:577874.
120. Rusz J, Krack P, Tripoliti E. From prodromal stages to clinical trials: the promise of digital speech biomarkers in Parkinson's disease. Neurosci Biobehav Rev 2024;167:105922.

Supporting Data

Additional Supporting Information may be found in the online version of this article at the publisher's web-site.