NBSE Evidence-Based Review: Test Summary Tables

Table of Contents

Contents

NBSE Evidence-Based Review: Test Summary Tables

Table of Contents

Attention

Test Name	Continuous Performance Test (CPT)
Description	A vigilance task requiring focusing of attention over relatively long time periods during which the patient detects and responds to target stimuli. It was originally developed by Beck et al (1956).[1] There are multiple versions of CPT with similar designs, but which use different modalities, stimuli, responses, and analyses.[1] The more commonly used versions are the Integrated Visual and Auditory CPT (IVA+Plus)[2, 3] Test of Variables of Attention (TOVA),[4] and the Conners CPT-II,[5] which are all computerized. These 3 versions can include a simple and a more complex task. In the simple task, letters of the alphabet are presented serially, and the task is to respond each time the letter is X (where X stands for a particular letter). In the more complex task, the patient responds only to the letter X if it is preceded by the letter Y (another letter). In other variations, the modality, the type of stimuli, and the way responses are analyzed may be altered.
Specific Functions Assessed	Sustained attention
Subscales	No subscales
Number of Items/Scoring	Number correct, omissions, commissions, reaction times, or d' and b derived from signal detection theory can all be scored.
Copyright Status	Copyrighted; IVA+Plus[2, 3] TOVA[4] and Conners CPT-II[5]
Administration Time	Variable depending on the version, typically 5-15 minutes
Normative Psychometric Data	This is available for some versions (Conners CPT-II, TOVA), predominantly for ages 7-21.[6, 7]
Sensitivity and Specificity	Vary across versions. Attention deficit hyperactivity disorder and traumatic brain injury[8] are the most studied populations, but the test has also been used in other disorders such as Tourette's syndrome[9] and pediatric cancer survivors.[10]
Advantages	It is a sensitive measure of sustained attention, which has been widely applied to various disorders (e.g., attention deficit disorder).
Limitations	Requires a computer and relatively long test time

NBSE Evidence-Based Review: Test Summary Tables

Table of Contents

References	1. Beck LH, Bransome ED, Jr., Mirsky AF, Rosvold HE, Sarason I. A continuous performance test of brain damage. J Consult Psychol 1956;20:343-350. 2. Sandford JA, Turner A. Integrated auditory and visual continuous performance test: Interpretation manual. Richmond, VA: BrainTrain, 1995. 3. Sandford JA, Turner A. Iva+Plus™: Integrated visual and auditory continuous performance test interpretation manual. Richmond, VA: BrainTrain, Inc., 2013. 4. Greenberg LM, Kindschi CL, Dupuy TR, Hughes SJ. Test of variables of attention continuous performance test: Clinical manual. Los Alamitos, CA: The TOVA Company, 2007. 5. Conners' continuous performance test II: Computer program for windows technical guide and software manual. North Tonawanda, NY: Multi-Health Systems, 2000. 6. Burton L, Pfaff D, Bolt N, et al. Effects of gender and personality on the conners continuous performance Test. J Clin Exp Neuropsychol 2010;32:66-70. 7. Conners CK, Epstein JN, Angold A, Klaric J. Continuous performance test performance in a normative epidemiological sample. J Abnorm Child Psychol 2003;31:555-562. 8. Tinius TP. The Integrated Visual and Auditory Continuous Performance Test as a neuropsychological measure. Arch Clin Neuropsychol 2003;18:439-454. 9. Shucard DW, Benedict RH, Tekok-Kilic A, Lichter DG. Slowed reaction time during a continuous performance test in children with Tourette's syndrome. Neuropsychology 1997;11:147-155. 10. de Ruiter MA, van Mourik R, Schouten-van Meeteren AY, Grootenhuis MA, Oosterlaan J. Neurocognitive consequences of a paediatric brain tumour and its treatment: A meta-analysis. Dev Med Child Neurol 2013;55:408-417.

NBSE Evidence-Based Review: Test Summary Tables

Table of Contents

Test Name	Corsi Blocks
Description	A test of visuospatial short-term working memory that utilizes blocks fixed on a board. The test was originally devised by Corsi[1, 2] as a non-verbal counterpart of the Hebb Digits Task. Patients view a board with nine 1 inch blocks irregularly distributed and fixed to a board. The blocks are numbered on the examiner's side, but the numbers cannot be seen by the patient. In the first part of the task, the patient's spatial span (maximum number of blocks that can be tapped in the correct sequence) is determined using a strategy similar to determining the digit span. Then 24 trials are performed using sequences 1 block more than the spatial span. Every third trial the sequence repeats, but all other trials are not repeated. Studies have differed in their use of recall order (forward vs. backward)[3, 4] and scoring criteria.[3] Standardization of the Corsi block-tapping test was outlined by Kessels et al.[4] Their description used a 9 block configuration and sequences derived from the reports of Smirni et al.[5] and Capitani et al.[6] Sequences were recalled in the forward order only. Up to 2 trials were performed for each block length, and testing was stopped when the patient failed to reproduce both sequences at a particular length.
Specific Functions Assessed	Non-verbal short term memory, spatial and temporal/sequential span and learning (if sequences are repeated)
Subscales	None when tested as suggested by Kessels et al.;[4] learning can be measured if repeating sequences are used[1, 2]
Number of Items/Scoring	The standard test[4] has 9 blocks and the number of trials depends on the Block Span (length of the last correctly repeated sequence, range 2-9). Sequence length is from 2 to 9 digits. Various scoring measures have been used. Corsi (1972) calculated a percentage based on the number of correct trials divided by the total number of trials to be learned.[1] The system suggested by Kessels et al (2000) calculated a Block Span, which was the length of the last correctly repeated sequence, and a Total Score, which was the product of the Block Span multiplied by the number of correct trials until the test was discontinued. The test was stopped when a patient failed to reproduce both trials at a given length.[4]
Copyright Status	Public Domain
Administration Time	Approximately 5-15 minutes, depending on the version used
Normative Psychometric Data	No comprehensive normative data; some normative data exists for healthy young adults ages 7-21[7] and older adults, age range 20-70[4]
Sensitivity and	It has been tested in a wide variety of populations, including normal adults of

Table of Contents

Specificity	various ages, patients with Alzheimer's disease, Korsakoff's syndrome, schizophrenia, learning disabilities and focal brain lesions following surgery or stroke.[3, 4]
Advantages	Testing is simple to perform, and may differentiate left vs. right hemisphere lesions. It has been tested in a wide variety of populations.
Limitations	Minimal normative data exists. There is limited of scoring standardization despite efforts.[4] Performance can be affected by changes in the test apparatus (particularly block placement) and by tapping sequence relative to the spatial location of the blocks. It may confound memory for the spatial location of the tapped block with memory for the sequential order of the blocks. Impaired performance is seen with both limitations in non-verbal memory and deficits in executive functions.
References	1. Corsi PM. Memory and the medial temporal region of the brain. 1972. 2. Milner B. Interhemispheric differences in the localization of psychological processes in man. Br Med Bull 1971;27:272-277. 3. Berch DB, Krikorian R, Huha EM. The Corsi block-tapping task: Methodological and theoretical considerations. Brain Cogn 1998;38:317-338. 4. Kessels RP, van Zandvoort MJ, Postma A, Kappelle LJ, de Haan EH. The Corsi block-tapping task: Standardization and normative data. Appl Neuropsychol 2000;7:252-258. 5. Smirni P, Villardita C, Zappala G. Influence of different paths on spatial memory performance in the Block-Tapping Test. J Clin Neuropsychol 1983;5:355-359. 6. Capitani E, Laiacona M, Ciceri E. Sex differences in spatial memory: A reanalysis of block tapping long-term memory according to the short-term memory level. Ital J Neurol Sci 1991;12:461-466. 7. Farrell Pagulayan K, Busch RM, Medina KL, Bartok JA, Krikorian R. Developmental normative data for the Corsi block-tapping task. J Clin Exp Neuropsychol 2006;28:1043-1052.

Test Name	Digit Span Backward
Description	A widely used test of auditory working memory and attentional capacity. The examiner presents a series of random digits (e.g., 3-6-5) usually at a rate of one digit per second and the patient is required to repeat these digits in the reverse order (e.g., 5-6-3). If the response is correct, the examiner then presents the next longer sequence of random digits (e.g., 4-6-2-9). If the patient fails to repeat the sequence correctly, the examiner presents a second trial. If the patient responds correctly, the next highest sequence is presented. This continues until the patient misses both trials or is able to repeat nine digits backward. Milberg et al. (1986) suggest giving the next longer sequence when a patient fails two trials because the digits are recalled correctly but in the wrong sequence.[1] The rationale is that being able to sequence digit names is not the same process as maintaining information in the correct sequence. The score for the patient's performance would not change; rather, it would be recorded as a clinical observation.
Specific Functions Assessed	Attention capacity, mental manipulation, working memory; it requires both receptive language and verbal repetition ability
Subscales	No subscales
Number of Items/Scoring	The score is the highest number of backward digits that can be accurately repeated.
Copyright Status	Public domain[1]
Administration Time	Typically 5 to 10 minutes
Normative Psychometric Data	Kaplan, et al. (1991) provide cumulative percentile data for Backward Digit spans of 6 or better are within normal limits; a span of 5 is marginal to normal limits, a span of 4 is definitely borderline, and a span of 3 is defective.[2] Monaco et al (2012), recently published Forward/Backward digit span results from 362 healthy adults, aged 20-80.[3] Digit Span tasks are a component of many different tests and norms are based on the specifics of the administration procedure [e.g., various editions of the *Wechsler Intelligence Scales (*Adult and Child*), the Wechsler Memory Scales (Adult and Child)*, the *Stanford-Binet*, and the *Clinical Evaluation of Language Fundamentals* (CELF)].
Sensitivity and Specificity	Vary across studies and patient populations. This test is much more sensitive than Digit Span Forward in patients with traumatic brain injury, early dementia, and right hemisphere dysfunction.
Advantages	A simple, rapidly administered bedside test that in combination with

[1] However, digit span tests and norms are components of numerous psychological tests (see below, under Normative Psychometric Data) as well as a number of computerized tests.

	Digit Span Forward serves as a good screening test for attention and working memory.
Limitations	Requires adequate hearing/speech. Examiner factors may affect performance (e.g., clarity of speech, speed of presentation). If malingering is suspected (suggested by digits backwards performance exceeding digits forward), additional strategies to determine reliability are required.
References	1. Milberg WP, Hebben N, Kaplan E. The Boston process approach to neuropsychological assessment. In: Grant I, Adams KM, eds. Neuropsychological assessment of neuropsychiatric disorder. New York: Oxford University Press, 1986: 65-86. 2. Kaplan E, Fein D, Morris R, Delis D. WAIS-R as a neuropsychological instrument. San Antonio, TX: Psychological Corporation, 1991. 3. Monaco M, Costa A, Caltagirone C, Carlesimo GA. Forward and backward span for verbal and visuo-spatial data: Standardization and normative data from an Italian adult population. Neurol Sci 2013;34:749-754.

Test Name	Digit Span Forward
Description	A simple test of auditory working memory. The digit span, as one form of span tests for attentional capacity, is probably the one in most common use for measuring the span of immediate verbal recall. The examiner presents a series of random digits (e.g., 3-6-5) usually at a rate of one digit per second (some tests vary the rate of presentation). The patient is required to repeat these digits exactly as presented. If the response is correct, the examiner then presents the next longer sequence of random digits (e.g., 4-6-2-9). If the patient fails to repeat the sequence correctly, the examiner presents a second trial. If the patient responds correctly, the next highest sequence is presented. This continues until the patient misses both trials or is able to repeat nine digits forward. The number of trials to "pass" a specific digit span length varies with the specific test, but is typically two.[1] Forward/backward digit tasks involve somewhat different neurocognitive processes and are affected in different ways by brain lesions and possibly aging[2].
Specific Functions Assessed	Attentional capacity, working memory, receptive language and verbal repetition ability
Subscales	No subscales
Number of Items/Scoring	The score is the highest number of forward digits that can be accurately repeated. Many psychological tests lump forward/backward digit spans into a single raw score and use that number to derive a Digit Span standard score.
Copyright Status	Public domain[2]
Administration Time	Typically 5 to 10 minutes
Normative Psychometric Data	Kaplan et al. suggest that it makes sense to utilize the raw score, given that a digit span of 5 to 8 was found in 89% of a large normative sample. (These authors also provide cumulative percentile data for both Forward and Backward Digits.) Forward Digit Spans of 6 or better are within normal limits; a span of 5 is marginal to normal limits, a span of 4 is definitely borderline, and a span of 3 is defective. Digit Span tasks are a component of many different tests and norms are based on the specifics of the administration procedure for the test [e.g., various editions of the *Wechsler* Intelligence *Scales (*Adult and Child*), the Wechsler Memory Scales (Adult and Child)*, the *Stanford-Binet*, the *Clinical Evaluation of Language Fundamentals* (CELF), and *the*

[2] However, digit span tests and norms are components of numerous psychological tests (see below, under Normative Psychometric Data).

	Comprehensive Test of Phonological Processing (CTOPP)]. There are also a number of computerized tests that incorporate digit span tests.
Sensitivity and Specificity	Vary across studies and patient populations.
Advantages	It is a simple, rapidly administered bedside test that serves as a good screening test for attentional disturbances.
Limitations	Requires adequate hearing/speech. Anxiety may significantly interfere with performance. Examiner factors may affect performance (e.g., clarity of speech, speed of presentation, etc.). If malingering is suspected (suggested by digits backwards exceeding digits forward), additional strategies to determine reliability are required.[3]
References	1. Lezak MD. Neuropsychological assessment, 3rd ed. New York: Oxford University Press, 1995. 2. Kaplan E, Fein D, Morris R, Delis D. WAIS-R as a neuropsychological instrument. San Antonio, TX: Psychological Corporation, 1991. 3. Heinly MT, Greve KW, Bianchini KJ, Love JM, Brennan A. WAIS digit span-based indicators of malingered neurocognitive dysfunction: Classification accuracy in traumatic brain injury. Assessment 2005;12:429-444.

Test Name	N-Back Test
Description	A complex tracking and continuous performance test of aspects of attention and working memory. It was first introduced by Kirchner in 1958[1] to examine age-differences in short-term memory retention. The patient is presented with a sequence of stimuli (e.g., auditory or visual). The task consists of indicating whether the current stimulus matches the one from *n* intervening stimuli (or steps) previously in the sequence. For example, using the numeric sequence "4-9-9-4-3-7-3," the 1-back task would match "9" to the immediately preceding "9"; a 2-back task matches "3" to "3", which occurred 2 digits previously, and the 3-back task matches "4" to "4", which occurred 3 digits previously. The difficulty can be adjusted by increasing the number of intervening stimuli *(n)* for the match. The difficulty can also be increased by the dual-task n-back task method, proposed by Jaeggi et al[2] in which two sequences are presented simultaneously, typically in different modalities (e.g., auditory and visual).
Specific Functions Assessed	Sustained and divided attention, working memory and mental flexibility
Subscales	No subscales
Number of Items/Scoring	Results can be reported as: correct responses, average correct response time, incorrect responses, average incorrect response time, and omission errors.
Copyright Status	Public Domain
Administration Time	Approximately 5 minutes
Normative Psychometric Data	Data are available for some versions. Normative data exist for healthy adolescents[3] and adults, age 21-80[4, 5] (note: these studies used different versions of the task, containing different visuospatial and/or verbal protocols).
Sensitivity and Specificity	Vary across studies. It has been used in patients with traumatic brain injury[6] and schizophrenia[7] among other disorders.
Advantages	A challenging task of working memory for which the difficulty can be adjusted. It has strong face validity and has been used widely as a measure of working memory in clinical and experimental settings.
Limitations	Questions have been raised about its construct validity since it has weak correlation with other tests of working memory. There is no standardized version that has robust normative data. Administration typically requires a computer.
References	1. Kirchner WK. Age differences in short-term retention of rapidly changing information. J Exp Psychol 1958;55:352-358. 2. Jaeggi SM, Seewer R, Nirkko AC, et al. Does excessive memory

	load attenuate activation in the prefrontal cortex? Load-dependent processing in single and dual tasks: Functional magnetic resonance imaging study. Neuroimage 2003;19:210-225. 3. Rigoli D, Piek JP, Kane R, Oosterlaan J. Motor coordination, working memory, and academic achievement in a normative adolescent sample: Testing a mediation model. Arch Clin Neuropsychol 2012;27:766-780. 4. Cansino S, Hernandez-Ramos E, Estrada-Manilla C, et al. The decline of verbal and visuospatial working memory across the adult life span. Age (Dordr) 2013. 5. Jaeggi SM, Buschkuehl M, Perrig WJ, Meier B. The concurrent validity of the N-back task as a working memory measure. Memory 2010;18:394-412. 6. Palacios EM, Sala-Llonch R, Junque C, et al. White matter integrity related to functional working memory networks in traumatic brain injury. Neurology 2012;78:852-860. 7. Ettinger U, Williams SC, Fannon D, et al. Functional magnetic resonance imaging of a parametric working memory task in schizophrenia: Relationship with performance and effects of antipsychotic treatment. Psychopharmacology (Berl) 2011;216:17-27.

Test Name	Oral Trail Making Test (oTMT)
Description	A short test of basic auditory attention and set-shifting. It is an oral version of the Trail Making Test (TMT) and was first described by Ricker and Axelrod in 1994.[1] The oTMT removes the visual and graphomotor components of the written TMT. As with the TMT, there are 2 parts: A and B. In part A, the patient counts out loud from 1 to 25 as quickly as possible. In part B, the patient is instructed to alternate between numbers and letters (e.g. 1-A-2-B-3-C) until he/she reaches 13. If the patient makes a mistakes on either task, they are directed back to the last correct item (for part A) or item pair (for part B) and must continue from there[2]. An almost identical task to oTMT-B is the alphanumeric sequencing test,[3, 4] which differs only in stopping at the letter L, instead of the number 13. In another variation, the Mental Alternation Task,[5, 6] patients first count from 1-20, then recite the alphabet and finally alternate between numbers and letters for 30 seconds.
Specific Functions Assessed	Basic auditory attention and sustaining minimal effort (Part A), cognitive set-shifting and basic executive control (Part B)
Subscales	Parts A and B are the subtests of oTMT
Number of Items/Scoring	oTMT-A has 25 numbers, and oTMT-B has 25 items. Scoring is based on the time to complete each of the parts. In the alphanumeric sequencing test, it is the time to complete 24 items. In the mental alternation task, only the alternation part is scored as the number of items completed in 30 seconds (maximum score is 52).
Copyright Status	Public domain
Administration Time	Less than 5 minutes
Normative Psychometric Data	There is limited normative data for several different types of patients. Ricker et al (1994) studied 3 patient groups, mean ages: 18.9, 31.9 and 83.5[1] while Mrazik et al., looked across the adult life span from 20-90 years.[7] Other groups have included those[1, 7] with cerebrovascular disease,[8] older patients with medical disease,[9] and a mixed clinical sample.[2]

Sensitivity and Specificity	Vary by clinical population; see normative data references. It has been studied in patients with stroke[8] and other medical illnesses.[2, 9]
Advantages	This is a simple test that requires no tools. oTMT-B performance is strongly correlated to written TMT-B performance. Some studies have suggested a fairly consistent written-to-oral TMT-B ratio around 2.5, but other studies have shown this to vary with age.[7, 10] The oTMT tests can be administered to patients with visual and motor impairments. When performed together with the written TMT, it may clarify the source of deficits.
Limitations	Patients who do not complete the task (i.e., are unable to get to 13 for part B) do not receive a score. Oral TMT-A is a fairly rote task and correlates variably with performance on the written TMT-A.[7] Oral TMT-B may be somewhat sensitive to aging and the effects of lower levels of education, potentially limiting its usefulness in these populations when patients are unable to complete the test.[9]
References	1. Ricker JH, Axelrod BN. Analysis of an Oral Paradigm for the Trail Making Test. Assessment 1994;1:47-52. 2. Abraham E, Axelrod B, Ricker JH. Application of the Oral Trail Making Test to a mixed clinical sample. Arch Clin Neuropsychol 1996;11:697-701. 3. Grigsby J, Kaye K. Alphanumeric sequencing and cognitive impairment among elderly persons. Percept Mot Skills 1995;80:732-734. 4. Grigsby J, Kaye K, Busenbark D. Alphanumeric sequencing: A report on a brief measure of information processing used among persons with multiple sclerosis. Percept Mot Skills 1994;78:883-887. 5. Jones BN, Teng EL, Folstein MF, Harrison KS. A new bedside test of cognition for patients with HIV infection. Ann Intern Med 1993;119:1001-1004. 6. McComb E, Tuokko H, Brewster P, et al. Mental alternation test: Administration mode, age, and practice effects. J Clin Exp Neuropsychol 2011;33:234-241. 7. Mrazik M, Millis S, Drane DL. The Oral Trail Making Test: Effects of age and concurrent validity. Arch Clin Neuropsychol 2010;25:236-243. 8. Ricker JH, Axelrod BN, Houtler BD. Clinical validation of the Oral Trail Making Test. Cogn Behav Neurol 1996;9:50-53. 9. Ruchinskas RA. Limitations of the Oral Trail Making Test in a mixed sample of older individuals. Clin Neuropsychol 2003;17:137-142. 10. Axelrod BN, Lamberty GJ. The Oral Trail Making Test. In: Poreh AA, ed. Neuropsychological assessment: A quantified process approach. Lisse, the Netherlands: Swets & Zeitlinger, 2006.

Test Name	Paced Auditory Serial-Addition Task (PASAT)
Description	A difficult test of sustained and divided auditory attention. It was first described by Gronwall[1, 2] as a method for evaluating recovery in patients with traumatic brain injury (TBI). The test has also been widely used in other conditions (e.g., clinical trials in multiple sclerosis).[3] In the test, a series of randomized numbers is presented. The patient listens to a series of numbers. Following the 2nd number, the patient responds with the sum of each consecutive pair of numbers. For example, if the series of numbers "4-3-8-2" is presented, then the patient answers "7" after "4-3," then "11" after "3-8," then "10" after "8-2." The digits are presented at four differing rates (e.g., 2.4, 2.0, 1.6 and 1.2 seconds). There are now 50- and 100-item short-form versions and computerized versions of the test.
Specific Functions Assessed	The test measures attention, working memory, auditory information processing speed, flexibility, and arithmetic abilities.
Subscales	No subscales
Number of Items/Scoring	The test score is the total number of correct sums given out of the total number possible for the particular version.
Copyright Status	Public Domain
Administration Time	10-15 minutes
Normative Psychometric Data	Normative data are available for several versions: for healthy adults ages 20-68 on the short forms (50 and 100-item),[4] and, healthy adults aged 17-40 on a computerized version.[5]
Sensitivity and Specificity	Vary across studies and clinical populations. It has been studied most extensively in patients with TBI[6] and MS.[3, 7, 8]
Advantages	One of only a few difficult attention/executive function tasks with considerable literature available for comparison, especially in TBI[6] and multiple sclerosis.[3, 7, 8] The PASAT test stimuli have been translated into 27 languages.
Limitations	There are significant practice effects over repeated measures. Typically requires audiotape or computer for presentation. Performance may also be influenced by math anxiety.
References	1. Gronwall DM. Paced auditory serial-addition task: A measure of recovery from concussion. Percept Mot Skills 1977;44:367-373. 2. Spreen O, Strauss E. A compendium of neuropsychological tests. New York: Oxford University press, 1998. 3. Cutter GR, Baier ML, Rudick RA, et al. Development of a multiple sclerosis functional composite as a clinical trial

	outcome measure. Brain 1999;122 (Pt 5):871-882. 4. Diehr MC, Cherner M, Wolfson TJ, et al. The 50 and 100-item short forms of the Paced Auditory Serial Addition Task (PASAT): Demographically corrected norms and comparisons with the full PASAT in normal and clinical samples. J Clin Exp Neuropsychol 2003;25:571-585. 5. Wingenfeld SA, Holdwick DJ, Jr., Davis JL, Hunter BB. Normative data on computerized paced auditory serial addition task performance. Clin Neuropsychol 1999;13:268-273. 6. Brenner LA, Terrio H, Homaifar BY, et al. Neuropsychological test performance in soldiers with blast-related mild TBI. Neuropsychology 2010;24:160-167. 7. Fischer JS, Rudick RA, Cutter GR, Reingold SC. The Multiple Sclerosis Functional Composite Measure (MSFC): An integrated approach to MS clinical outcome assessment. National MS society clinical outcomes assessment task force. Mult Scler 1999;5:244-250. 8. Rudick R, Antel J, Confavreux C, et al. Recommendations from the national multiple sclerosis society clinical outcomes assessment task force. Ann Neurol 1997;42:379-382.

Test Name	Sequential Operations Series (SOS)
Description	Brief, separate tests that examine similar functions of basic attention and working memory: Serial 7 subtraction (SSS) As originally described,[1] a patient is asked to take 7 away from 100, and then to take 7 away from each answer obtained. Up to 2 minutes was given for the test. Each subtraction is considered its own unit, so mistakes on one calculation do not bias the others. Other versions of the test score only the first 5 subtractions.[2] WORLD backwards (WB) (or similar words) The patient is asked to spell WORLD backwards. Some versions of the Mini-Mental State Examination (MMSE) have patients spell WORLD forwards first. Alphabet backwards (AB) The patient is asked to recite the alphabet in reverse. After any error the correct letter is given and the patient prompted to continue from that point.[3] Months of the year backwards (MOYB) In some versions the individual is asked to recite the months forwards starting from January, and then backwards starting from December; other versions just use the backwards task.
Specific Functions Assessed	Simple attention, working memory; also reflects information processing speed,[3, 4] some frontal lobe functions, like mental control,[5,6] and intellectual efficiency.[1]
Subscales	No subscales
Number of Items/Scoring	Serial 7 Subtraction (SSS) Up to 14 subtractions are possible. As initially described, scoring was based on a normalized response index (number of correct responses / number of total responses) x 14. Errors were noted to take on a variety of forms including different patterns of repeating the terminal digit.[1] The number of seconds to complete the task can also be measured.[3] In a short form of the task, used on the MMSE,[2] only 5 items are assessed and the score is based on the number of correct responses. WORLD backwards (WB) Usually scored by giving 1 point for each letter in the correct ordinal position. However, this does not account for difference in the types of errors such as omissions in the middle of the word (DLRW – scored 3/5) vs. those at the beginning (LROW – scored 0/5). Alphabet backwards (AB) Scored as the number of seconds to complete the task, the number of errors made, and a performance index, which is the number of seconds to complete the task / the number of correct responses.[3]

	Months of the year backwards (MOYB) Various scoring rules: The score is sometimes just the time taken, with errors ignored. On the WMS-IV, the score is a combination of time taken and number of errors.[7] [8]
Copyright Status	Public domain[3]
Administration Time	Less than 3 min per task
Normative Psychometric Data	A few studies have examined individual SOS tests in clinical populations. These include SSS in children ages 8-15,[1] adults ages 18-63[9] and greater than 65,[10] patients with cardiac disease,[3] high school athletes,[4] and patients with frontal lesions.[6] Moore et al. examined the degree to which subtracting other numbers would be equivalent to SSS and found that 6, 8 and 9 were roughly equivalent to 7.[11] MOYB has been studied in young athletes,[4] adults, ages 18-94,[8, 12] frontal lesion patients [6] and patients with delirium.[13] AB has been used in studies of adults over 65[10] and patients with cardiac disease.[3] Finally, WB was used in testing of frontal patients [6] and adults over 65.[10]
Sensitivity and Specificity	Varies both by test and within test by scoring methodology. For SSS and AB the number of seconds to complete the test was more sensitive than the number of errors in distinguishing cardiac patients from controls.[3] Because many healthy, normal individuals have difficulty with SSS, the test can be less specific than the others listed, as reduced performance may not indicate a true decline from the patient's baseline.[4] SSS has been studied in various psychiatric patients, but did not clearly distinguish the patient groups.[1]
Advantages	Rapidity and ease of administration of all the tests
Limitations	SSS is sensitive to age, education and gender. A high proportion of normal, highly educated adults make errors on SSS (50% with 16 or more years of education), and only 42% could perform all subtractions.[9] There is poor correlation of performance between SSS and WB,[10] and SSS and MOYB.[4] Only 2/3 of normal high school students overall could recite MOYB without error, and women in their later teens (16-19) were better than men (84.4% vs. 70.4%),[14] raising the question of reduced specificity of the error measure for this test in younger age groups.
References	1. Hayman M. Two minute clinical test for measurement of intellectual impairment in psychiatric disorders. Arch Neurol Psychiatry 1942;47:454-464. 2. Folstein MF, Folstein SE, McHugh PR. "Mini-mental state". A practical method for grading the cognitive state of patients for the clinician. J Psychiatr Res 1975;12:189-198. 3. Williams MA, LaMarche JA, Alexander RW, Stanford LD, Fielstein

[3] Some are also included in copyrighted test batteries, like SSS and WB in the MMSE [2] and MOYB in the Wechsler Memory Scale-IV and Brief Cognitive Status Exam[7].

EM, Boll TJ. Serial 7s and alphabet backwards as brief measures of information processing speed. Arch Clin Neuropsychol 1996;11:651-659.
4. Young CC, Jacobs BA, Clavette K, Mark DH, Guse CE. Serial sevens: Not the most effective test of mental status in high school athletes. Clin J Sport Med 1997;7:196-198.
5. Luria AR. Higher cortical functions in man. New York: Basic Books, 1966.
6. Ettlin TM, Kischka U, Beckson M, Gaggiotti M, Rauchfleisch U, Benson DF. The frontal lobe score: Part I: Construction of a mental status of frontal systems. Clin Rehabil 2000;14:260-271.
7. Wechsler D. Wechsler Memory Scale - IV. San Antonio, TX: Pearson, 2009.
8. Ball LJ, Bisher GB, Birge SJ. A simple test of central processing speed: An extension of the short blessed Test. J Am Geriatr Soc 1999;47:1359-1363.
9. Smith A. The serial sevens subtraction test. Arch Neurol 1967;17:78-80.
10. Ganguli M, Ratcliff G, Huff FJ, et al. Serial sevens versus world backwards: A comparison of the two measures of attention from the MMSE. J Geriatr Psychiatry Neurol 1990;3:203-207.
11. Moore PN, Pierce D, Graybill D. Digits of equivalent difficulty in the serial subtraction test. Percept Mot Skills 1980;50:940-942.
12. Ostberg P, Hansson V, Haagg S. Adult norms and test-retest reliability for the Months Backward test: Durational and response accuracy measures. Logoped Phoniatr Vocol 2012;37:11-17.
13. Farrell Pagulayan K, Busch RM, Medina KL, Bartok JA, Krikorian R. Developmental normative data for the Corsi block-tapping task. J Clin Exp Neuropsychol 2006;28:1043-1052.
14. Jinguji TM, Bompadre V, Harmon KG, et al. Sport concussion assessment tool-2: Baseline values for high school athletes. Br J Sports Med 2012;46:365-370.

Test Name	Stroop Test
Description	A widely used test of selective attention and cognitive flexibility. The test was first developed by John Stroop (1935) to study a previously described interference effect that occurs during normal visual-verbal processing.[1] It has since been used widely to measure attention and mental flexibility in many neurologic disorders. Performance on the test is also strongly related to IQ scores and can improve with practice. The original version has four parts. It includes three white cards, each containing 10 rows of five items. In Part 1, the patient reads color names printed in black type. In Part 2, the patient reads color names where the color of the print and the word are different. In Part 3, the patient names colors of a series of different colored squares. In Part 4, the patient is given the card used in Part 2 and asked to name the color of the ink in which the word is printed, rather than reading the name of the color. If the patient makes a mistake on any of the parts, the examiner tells them "No" and they have to stop and correct the error.[2] The outcome variable is the difference in color-naming speed. Patients must inhibit their pre-potent response (i.e. reading the printed word) and instead must attend to and employ a different naming strategy (i.e. saying the name of the color the word is printed in). Cognitive flexibility is also required and is enhanced by requiring the patient to shift from reading the name of the color in which the word is printed to naming the color. The difference between the speed of naming of Part 4 and Part 3 (Part 4 minus Part 3) is called the "color-word interference effect." Working memory is also involved in the interference effect in that it requires that the test-taker hold the goal in mind for the duration of the task. Another variation of the task, called the emotional Stroop, uses emotional words printed in different colors.[3] Multiple studies of patients with various types of psychopathology have shown slower color naming for the emotional vs. neutral words, for example, slower processing in patients with depression vs. controls of negative vs. neutral words.[3]
Specific Functions Assessed	Visual selective attention, visual scanning, motor speed, working memory, cognitive flexibility
Subscales	Vary with the version of the task. The Delis-Kaplan Executive Function System (D-KEFS) Color-Interference test has four subtests: Color-Naming, Color-Word Naming, Interference, and Inhibition/Switching. Normative contrast scores compare performance across these various conditions. Norms are also provided for number of errors, and

	corrected and uncorrected errors.
Number of Items/Scoring	Varies with specific version of task (range 24 to 100 items). Scoring for each trial type is based on the number of correct responses in a fixed amount of time, typically within 45 seconds. There is an "interference score" which reflects the color-word interference effect (see Description above). Note that there is only moderate correlation between scores on different versions of the Stroop. The Emotional Stroop task does not have a standardized version, but detailed examples can be found in various research reports, including work by Smith et al.[4]
Copyright Status	Copyrighted[4] and public domain[5]
Administration Time	5-10 minutes
Normative Psychometric Data	Many versions of the standard Stroop test have normative data available, but it may be restricted to a limited age range. The Victoria version has norms for ages 18-94.[5] The D-KEFS version has data available for ages 8 to 89 years.[6] Normative data on shorter versions of the D-KEFS in older adults has also been reported recently.[7]
Sensitivity and Specificity	Limited sensitivity and specificity varies with population being tested. Longer response times and/or increased errors on the interference subtest has been seen in several patient groups including Alzheimer's disease,[8, 9] Parkinson's disease,[10] and multiple sclerosis.[11] and for the emotional Stroop in a variety of psychopathologies.[3]
Advantages	Sensitive measure of sustained and controlled attention, which has been widely applied to various disorders.
Limitations	Various versions are limited in that norms are restricted to certain age groups, and some versions lack norms for error scores. It is important to supplement the Interference score with other assessments. The time differential represented by the Interference score reflects a combination of tests such that some patients have average-range time scores but at the expense of generating many errors. Test performance is affected by visual disturbance (e.g., low contrast sensitivity) and certain types of color-blindness, anomia or slow naming speed (e.g., dyslexia). There is also a time-of-day effect in older adults.

[4] Two commonly used English versions of the Stroop, the Golden Stroop Color and Word Test and D-KEFS, are copyrighted
[5] The Victoria version is in the public domain and users may make their own stimuli.

References	1. Stroop JR. Studies of interference in serial verbal reactions. J Exp Psychol 1935;18:643-662. 2. Killian G. The Stroop color-word interference test. In: D K, R S, eds. Test critiques. Kansas City, MO: Test Corporation of America, 1985: 751-758. 3. Williams MA, LaMarche JA, Alexander RW, Stanford LD, Fielstein EM, Boll TJ. Serial 7s and alphabet backwards as brief measures of information processing speed. Arch Clin Neuropsychol 1996;11:651-659. 4. Smith P, Waterman M. Processing bias for aggression words in forensic and nonforensic samples. Cogn Emot 2003;17:681-701. 5. Troyer AK, Leach L, Strauss E. Aging and response inhibition: Normative data for the Victoria Stroop Test. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2006;13:20-35. 6. Delis DC. Delis-Kaplan Executive Function System (D-KEFS). San Antonio, TX: The Psychological Corporation, 2001. 7. LaMarre AK, Rascovsky K, Bostrom A, et al. Interrater reliability of the new criteria for behavioral variant frontotemporal dementia. Neurology 2013;80:1973-1977. 8. Hutchison KA, Balota DA, Duchek JM. The utility of Stroop task switching as a marker for early-stage Alzheimer's disease. Psychol Aging 2010;25:545-559. 9. Balota DA, Tse CS, Hutchison KA, Spieler DH, Duchek JM, Morris JC. Predicting conversion to dementia of the Alzheimer's type in a healthy control sample: The power of errors in Stroop color naming. Psychol Aging 2010;25:208-218. 10. Obeso I, Wilkinson L, Casabona E, et al. Deficits in inhibitory control and conflict resolution on cognitive and motor tasks in Parkinson's disease. Exp Brain Res 2011;212:371-384. 11. Lynch SG, Dickerson KJ, Denney DR. Evaluating processing speed in multiple sclerosis: A comparison of two rapid serial processing measures. Clin Neuropsychol 2010;24:963-976.

Test Name	Trail-Making Test
Description	A brief, 2-part test of basic attention and working memory. The original Trail-Making Test (TMT) was developed by a team of U.S. Army psychologists as part of the Army Individual Test Battery (1944).[1] Different administration and scoring systems were subsequently introduced. It has developed into a widely used instrument, with versions incorporated in to the Delis-Kaplan Executive Function System (D-KEFS)[2] and the Halstead-Reitan Battery (HRB).[3] In Part A (TMT-A), the patient draws lines connecting consecutively numbered circles. In Part B (TMT-B), the patient draws lines alternating between circles containing numbers and circles containing letters, fulfilling a specific, alternating sequence (i.e., 1-A, 2-B, 3-C). The measured value is the time to completion for each part. In some adapted versions, like the HRB, the examiner notes errors and points them out to the patient during the test. Other factors that might affect performance are visual scanning and motor speed.
Specific Functions Assessed	Attention/visual scanning, motor speed, working memory
Subscales	Of the common versions, the D-KEFS involves five subscales: 1. *Visual Scanning*. This is a cancellation task in which the patient is asked to identify and cross out a specific target on an 11”x17” page. This test would also serve to identify hemispatial neglect. Omission/commission errors are included in the scoring as well as time to complete. 2. *Number Sequencing.* This is a number sequencing task (similar to TMT-A) 3. *Letter Sequencing.* This involves connecting letters of the alphabet in sequence. 4. *Number-Letter Switching.* The patient is asked to draw lines, switching between numbers and letters in sequence (e.g., 1-A, 2-B, etc.). This is the primary executive function task, similar to TMT-B. 5. *Motor Speed.* The patient traces over a line connecting a series of circles.
Number of Items/Scoring	On the two-part TMT, the times to completion for A and B are recorded and the difference of time-to-completion on B minus A is used as the variable and compared to norms. A large B minus A difference reflects attentional difficulties and/or problems with working memory. An alternative computation involves the ratio B/A, but this is felt to be less reliable.[4]

Attention

Copyright Status	Copyrighted[6] and public domain[7]
Administration Time	Variable depending on the version and type of patient, typically 5-20 minutes
Normative Psychometric Data	Norms are specific to the version. Ample norms are available for ages 18-89.[5] The D-KEFS version provides age-based norms for each of the 5 subtests (listed above). Norms for ages 8 to 89 years are provided for each subscale task and error analysis. In addition to time-completion scores, error scores for sequencing errors versus set-loss error scores are provided. Contrasts between performances on the various subscale tasks (with norms) are also provided.
Sensitivity and Specificity	Vary across versions. Performance on the Trail Making Test, particularly part B, correlates with reduced driving performance[6, 7] and mobility impairment with aging.[8] The TMT has also been used in studies of Huntington's disease,[9] and traumatic brain injury[10] among other disorders.
Advantages	Sensitive measure of sustained attention, attentional control, and set-shifting, which has been widely applied to various disorders (e.g., attention deficit disorder). Careful observation of the way patients approaches the task, the type of errors they make, and analysis of the type errors contribute to the test's value.
Limitations	Task performance is affected by a number of factors (e.g., age, educational status) in addition to neurocognitive issues. These need to be factored in to the analysis.[11, 12]
References	1. Army individual Test battery. In: War Department, ed. Manual of Directions and Scoring. Washington, DC: Adjutant General's Office, 1944. 2. Delis DC. Delis-Kaplan Executive Function System (D-KEFS). San Antonio, TX: The Psychological Corporation, 2001. 3. Lezak MD. Neuropsychological assessment, 3rd ed. New York: Oxford University Press, 1995. 4. Martin TA, Hoffman NM, Donders J. Clinical utility of the Trail Making Test ratio score. Appl Neuropsychol 2003;10:163-169. 5. Tombaugh TN. Trail Making Test A and B: Normative data stratified by age and education. Arch Clin Neuropsychol 2004;19:203-214. 6. Emerson JL, Johnson AM, Dawson JD, Uc EY, Anderson SW, Rizzo M. Predictors of driving outcomes in advancing age. Psychol Aging 2012;27:550-559. 7. Hargrave DD, Nupp JM, Erickson RJ. Two brief measures of executive function in the prediction of driving ability after acquired brain injury. Neuropsychol Rehabil 2012;22:489-500.

[6] The D-KEFS is copyrighted and commercially available.
[7] The original Trail-Making Test and Halstead-Reitan adapted version are in the public domain.

	8. Vazzana R, Bandinelli S, Lauretani F, et al. Trail Making Test predicts physical impairment and mortality in older persons. J Am Geriatr Soc 2010;58:719-723. 9. O'Rourke JJ, Beglinger LJ, Smith MM, et al. The Trail Making Test in prodromal Huntington disease: Contributions of disease progression to test performance. J Clin Exp Neuropsychol 2011;33:567-579. 10. Perianez JA, Rios-Lago M, Rodriguez-Sanchez JM, et al. Trail Making Test in traumatic brain injury, schizophrenia, and normal ageing: Sample comparisons and normative data. Arch Clin Neuropsychol 2007;22:433-447. 11. Bornstein RA, Suga LJ. Educational level and neuropsychological performance in healthy elderly subjects. Dev Neuropsychol 1988;4:17-22. 12. Stuss DT, Stethem LL, Pelchat G. Three tests of attention and rapid information processing: An extension. Clin Neuropsychol 1988;2:246-250.

Executive Functioning

Test Name	Anti-Saccade Test
Description	A simple test which measures the degree of volitional control of behavior and inhibition of automatic, reflexive behaviors. It specifically tests the control of eye movements, or saccades, as a representative of other forms of volitional behavior.[1] "Anti-saccade" refers to the intentional directing of eye movements away from a visual stimulus. In contrast, "pro-saccade" refers to the automatic (or reflexive) behavior of orienting gaze toward a visual stimulus. In its simplest form, the test involves only a patient and an examiner. The examiner positions herself in front of the patient and instructs the patient first to fixate on a central point. The examiner than presents a peripheral, visual stimulus (often an extended finger) in the patients' visual field. The patient is instructed to "look away" ("eyes-opposite") from the side of the presented stimulus. Performing the anti-saccade requires inhibition of a "pre-potent" response to look toward the stimulus. The test can also be done with presentation software on a computer screen, also utilizing eye-tracking hardware to record latency times.
Specifics Functions Assessed	Inhibition of automatic behavior, volitional control of movement/action, flexible control over behavior
Subscales	No subscales
Number of Items/Scoring	Pro-saccade errors are estimated to happen in 10-20% of trials in young healthy controls; errors increase with age and improve with practice.[2] The test can be scored as a simple error rate over 5-10 trials. Latency for both error and correct responses can also be measured with eye-tracking hardware (e.g., Eye Link system©). Anti-saccade results are often compared to performance on the "pro-saccade" (control) test.
Public Domain/Copyright Status	Public domain
Administration Time	3 minutes
Normative Psychometric Data	Normative data exists for healthy controls ages 18-25, 70s, 80s.[3]
Sensitivity and Specificity	It has been studied in many disorders, including schizophrenia[4], ADHD, AD, PD, Tourette disorder, healthy controls, bipolar disorder, PSP, and FTD.[4, 5]

Advantages	The simplest, non-computerized version is an easy to learn, easy to administer, "bed-side" assessment. It has been validated as effective to detect executive dysfunction in multiple populations.[6] Anti-saccade errors have been used with reliability to study progression of AD.[5]
Limitations	It is not sensitive in detecting early impairment from degenerative disease. There is a normal decay of saccade initiation speed and pro-saccade errors with aging.
References	1. Munoz DP, Everling S. Look away: The anti-saccade task and the voluntary control of eye movement. Nat Rev Neurosci 2004;5:218-228. 2. Evdokimidis I, Smyrnis N, Constantinidis TS, et al. The antisaccade task in a sample of 2,006 young men. I. Normal population characteristics. Exp Brain Res 2002;147:45-52. 3. Klein C, Fischer B, Hartnegg K, Heiss WH, Roth M. Optomotor and neuropsychological performance in old age. Exp Brain Res 2000;135:141-154. 4. Gooding DC, Tallent KA. The association between antisaccade task and working memory task performance in schizophrenia and bipolar disorder. J Nerv Ment Dis 2001;189:8-16. 5. Crawford TJ, Higham S, Renvoize T, et al. Inhibitory control of saccadic eye movements and cognitive impairment in Alzheimer's disease. Biol Psychiatry 2005;57:1052-1060. 6. Hellmuth J, Mirsky J, Heuer HW, et al. Multicenter validation of a bedside antisaccade task as a measure of executive function. Neurology 2012;78:1824-1831.

Test Name	Frontal Assessment Battery (FAB)
Description	A brief, bedside battery that assesses the presence and severity of executive dysfunction. It involves testing features of both cognitive and motor functions. The FAB was initially designed in 2000 as a bedside tool help differentiate patients with frontotemporal dementia from Alzheimer disease dementia.[1] The battery consists of tests designed to detect and measure prehension behavior, verbal fluency, the ability to execute and control basic motor behaviors, and the ability to carry out low-level abstract thinking. It has also been studied as a tool to assess progression of executive impairment through the course of fronto-temporal dementia.[2]
Specific Function Assessed	Motor programming, conceptualization (abstraction), mental flexibility, sensitivity to interference, inhibitory control, and environmental autonomy
Subscales	6 subscales
Number of Items/Scoring	Each subscale is worth up to 3 points; the maximum score is 18 points
Public Domain/Copyright Status	Public Domain
Administration Time	10 minutes
Normative Psychometric Data	Available primarily for ages 60-80.[3] Some data exists for "old-old" groups (> 80 years old)[4, 5] and there more limited data is available for ages 20 to 60.[5] There is considerable variability based on education, with lower mean scores accepted as "normal" for those with less years of education.
Sensitivity and Specificity	Vary across studies. Patient populations that have been studied include MCI, AD dementia, ALS, MCI, PD, and substance abuse disorders.[3, 6-12]
Advantages	It is brief, easy to administer, well-tolerated by patients, and assesses several different functions mediated by frontal lobe networks. It may be used as a global screen for various aspects of executive dysfunction.
Limitations	Although several studies of elderly adults yield a similar range of scores, there are no clearly established cut-off scores. Cohorts with healthy controls often have less than 100 subjects, limiting the power of normative values. Performance on the FAB appears to be strongly influenced by education, potentially confounding results.

References	1. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: A Frontal Assessment Battery at bedside. Neurology 2000;55:1621-1626. 2. Gordon E, Rohrer JD, Kim LG, et al. Measuring disease progression in frontotemporal lobar degeneration: A clinical and MRI study. Neurology 2010;74:666-673. 3. Ahn SW, Kim SH, Kim JE, et al. Frontal assessment battery to evaluate frontal lobe dysfunction in als patients. Can J Neurol Sci 2011;38:242-246. 4. Iavarone A, Lore E, De Falco C, et al. Dysexecutive performance of healthy oldest old subjects on the Frontal Assessment Battery. Aging Clin Exp Res 2011;23:351-356. 5. Appollonio I, Leone M, Isella V, et al. The Frontal Assessment Battery (FAB): Normative values in an Italian population sample. Neurol Sci 2005;26:108-116. 6. Chong MS, Lim WS, Chan SP, et al. Diagnostic performance of the Chinese Frontal Assessment Battery in early cognitive impairment in an Asian population. Dement Geriatr Cogn Disord 2010;30:525-532. 7. Cunha PJ, Nicastri S, de Andrade AG, Bolla KI. The frontal assessment battery (FAB) reveals neurocognitive dysfunction in substance-dependent individuals in distinct executive domains: Abstract reasoning, motor programming, and cognitive flexibility. Addict Behav 2010;35:875-881. 8. Domingues SC, Mendonca JB, Laranjeira R, Nakamura-Palacios EM. Drinking and driving: A decrease in executive frontal functions in young drivers with high blood alcohol concentration. Alcohol 2009;43:657-664. 9. Kaszas B, Kovacs N, Balas I, et al. Sensitivity and specificity of Addenbrooke's Cognitive Examination, Mattis Dementia Rating Scale, Frontal Assessment Battery and Mini Mental State Examination for diagnosing dementia in Parkinson's disease. Parkinsonism Relat Disord 2012;18:553-556. 10. Morimoto N, Kurata T, Sato K, Ikeda Y, Sato S, Abe K. Frontal dysfunctions of ALS-PBP patients in relation to their bulbar symptoms and rCBF decline. J Neurol Sci 2012;319:96-101. 11. Oshima E, Terada S, Sato S, et al. Frontal assessment battery and brain perfusion imaging in Alzheimer's disease. Int Psychogeriatr 2012;24:994-1001. 12. Tessitore A, Amboni M, Esposito F, et al. Resting-state brain connectivity in patients with Parkinson's disease and freezing of gait. Parkinsonism Relat Disord 2012;18:781-787.

Test Name	Frontal Behavioral Inventory (FBI)
Description	A brief questionnaire, administered to caregivers, that is designed to detect a range of behavioral and personality changes, most often seen in fronto-temporal disorders. The Frontal Behavioral Inventory (FBI) was developed and standardized in order to differentiate the behavioral variant of Fronto-temporal Dementia (bvFTD) from other dementias, such as Alzheimer's Disease and vascular dementia.[1] It has since been used to measure the severity of behavioral problems in other degenerative diseases as well as conditions that can affect frontal lobe networks, like stroke and traumatic brain injury.[2]
Specific Functions Assessed	Negative ("deficit") behaviors, like apathy, aspontaneity, indifference, inflexibility, personal neglect, disorganization, loss of insight, loss of comprehension and alien hand syndrome; Positive ("disinhibited") behaviors, like perseveration, obsessiveness, irritability, excessive jocularity, social inappropriateness, impulsivity, utilization behavior, and incontinence.
Subscales	2 subscales with 12 items each
Number of Items/Scoring	It consists of 24 questions. Each question is worth 0 to 3 points. A maximum score is 72 points. Scores above 40 indicate "Severe" disease, 30 to 39 "Moderate", and 25 and 29 "Mild".
Copyright Status	Public Domain
Administration Time	20 minutes
Normative Psychometric Data	None available. Demographic factors do not influence performance.
Sensitivity and Specificity	Varies across studies and depends on cut-off score used. Some have shown a 97% sensitivity and 95% specificity in discriminating bvFTD from other degenerative dementias with a cut-off score of 23.[2] Patient populations that have been studied include AD, large ischemic strokes, PSP, FTD, VaD, ALS, and TBI.[3-5]
Advantages	It is easy to administer. It captures presence or absence and severity of various behavioral symptoms from caregivers. It complements the history and adds specific severity ratings.
Limitations	A skilled interviewer is necessary to ensure that the questions are understood by the caregiver. The FBI has to be administered without the patient present.

References	1. Kertesz A, Davidson W, Fox H. Frontal behavioral inventory: Diagnostic criteria for frontal lobe dementia. Can J Neurol Sci 1997;24:29-36. 2. Milan G, Lamenza F, Iavarone A, et al. Frontal behavioural inventory in the differential diagnosis of dementia. Acta Neurol Scand 2008;117:260-265. 3. Kertesz A, Nadkarni N, Davidson W, Thomas AW. The frontal behavioral inventory in the differential diagnosis of frontotemporal dementia. J Int Neuropsychol Soc 2000;6:460-468. 4. Gordon PH, Wang Y, Doorish C, et al. A screening assessment of cognitive impairment in patients with ALS. Amyotroph Lateral Scler Other Motor Neuron Disord 2007;8:362-365. 5. Pachalska M, Talar J, Kurzbauer H, Franczuk B, Grochmal-Bach B, Macqueen BD. Differential diagnosis of frontal syndrome in patients with closed-head injuries. Ortopedia, Traumatologia, Rehabilitacja 2002;4:81-87.

Test Name	Luria Hand-Sequencing Test
Description	A brief, sequential motor task, designed to assess execution of a learned motor program, inhibitory control, and working memory/attentional flexibility. It can be done as an isolated "bedside" test but it is also contained within many larger test batteries, such as the Frontal Assessment Battery (FAB) and the Executive Interview (EXIT).[1] It was developed by Russian neuropsychologist Alexander Luria in the middle of the 20th century.[2] One of Luria's observations was that patients with substantial frontal lobe lesions were unable to alter their responses to a motor task when the additional steps were added or the order of steps was changed.[3] In the hand-sequencing test, patients are usually shown a series of 3 different hand-positions by the examiner (e.g. right fist in the left palm, then palm against palm). They are then asked to repeat these hand-positions, in order, on their own. Different sequences can be used sequentially. Patients are allowed to observe and learn the different positions before performing the test. Those with frontal lobe lesions or frontal network disruption often have difficulty keeping the simple hand-positions and their order in mind to repeat the sequence.[4]
Specific Functions Assessed	Different aspects of executive function, including execution of a learned motor program, inhibitory control, attentional flexibility, working memory, and motor planning
Subscales	No subscales
Number of Items/Scoring	Scores are based upon observation of the patient performing the series of different hand-positions. Healthy people should be able to learn and then execute the correct sequence of hand-positions at least 3 times, without error. Different test batteries have different scoring schemes. In the FAB, there is a 3 point-scale, where patients earn 3 points if they carry out the correct sequence 6 times, 2 points if they carry it out 3 times, point if they are unable to perform the series alone but can do so along with the examiner, and 0 points if the patient cannot perform the hand sequence series even with the examiner.
Copyright Status	Public Domain
Administration Time	Approximately 2-3 minutes
Normative Psychometric Data	Little normative data exists for the test alone. Ample normative data exist for the larger batteries (FAB, EXIT) that this test is embedded within.
Sensitivity and Specificity	Vary across studies. Patient populations that have been studied include dementias such as MCI, AD, and bvFTD, patients with traumatic brain injury and autism, and patients with psychiatric disorders, such as

	schizophrenia and ADHD spectrum.[5, 6] There is good inter-rater reliability and internal consistency is suggested based upon the FAB.[1]
Advantages	Easy to administer, brief, well-tolerated by patients, and sensitive to cultural differences. Provides brief, gross look at frontal lobe function.
Limitations	Performance may be hindered by other compounding factors such as disability, weakness, medication intake, sensory limitation, or other medical disease which may affect the joints, muscles, or bones. Also, there is limited normative data, and it is not a comprehensive test by itself.
References	1. Dubois B, Slachevsky A, Litvan I, Pillon B. The FAB: A Frontal Assessment Battery at bedside. Neurology 2000;55:1621-1626. 2. Christensen A-L. Luria's neuropsychological investigation : Text. Copenhagen: Munksgaard, 1979. 3. Pribram KH. Editorial foreword. In: Luria AR, ed. The working brain: An introduction to neuropsychology. New York: Basic Books, 1973. 4. Luria AR. Higher cortical functions in man. New York: Basic Books, 1980. 5. Weiner MF, Hynan LS, Rossetti H, Falkowski J. Luria's three-step test: What is it and what does it tell us? Int Psychogeriatr 2011;23:1602-1606. 6. Sander RD. Motor examinations in psychiatry. Psychiatry (Edgmont) 2010;7:37-41.

Test Name	Proverb Interpretation Tests
Description	A test that evaluates figurative/conceptual thinking. It is designed to detect the ability to interpret meaning beyond a concrete or literal explanation. In the modern literature, Gorham[1] reported using proverbs as early as the 1950s to understand abstract thinking in patients with schizophrenia. Typically, patients are asked to interpret the meaning of proverbs (e.g. "All that glitters is not gold"), with or without an associated pictorial representation. Proverbs have been used as a diagnostic tool across a variety of disciplines, including psychiatry, psychology, neurology, and speech-language pathology. The most widely used current proverbs set is that which constitutes the Proverbs Sub-Test of the Delis-Kaplan Executive Function System (D-KEFS), which is modeled after Gorham's original instrument. Another commonly used set is the 10-item set which constitutes the proverb interpretation test of Barth and Kufferle.[2]
Specifics Functions Assessed	Fundamental verbal skills, concrete (literal) interpretation, ability to disregard/inhibit concrete meaning for abstract meaning, higher-level integration of individual words or ideas, understanding metaphor
Subscales	No subscales
Number of Items/Scoring	There are several different commonly used proverb sets. Typically, 4-8 proverbs are chosen. Tests can include 2 conditions: 1) a free response condition, 2) a multiple choice condition. Scoring can simply be the number of proverbs felt to be accurately interpreted by the examiner. The D-KEFS Proverbs Sub-test (8 proverbs) is the only commonly used set with a numerical, graded system, which uses a 4-point scale for each interpretation (called the "Achievement Score").
Public Domain/Copyright Status	Different proverb sets/tests : 1. Delis-Kaplan Executive Function System (D-KEFS)-Proverbs Sub-test: Copyrighted 2. The Proverbs Test (Gorham, 1956)[1]: Public Domain 3. Proverb interpretation test of Barth and Kufferle[2]: Public Domain 4. The Familial and Novel Language Comprehension Test (FANL-C) [Verbal and non-verbal idioms]: Public Domain (http://blog.emerson.edu/daniel_kempler/fanlc.html)
Administration Time	10-12 minutes

Normative Psychometric Data	Little or no normative data exists for most proverb tests/sets, including the Proverbs Test (Gorham), FANL-C, and the proverb interpretation test (Barth and Kufferle). There is ample normative data for the D-KEFS Proverbs Sub-Test for ages 8-89.[3]
Sensitivity and Specificity	Populations that have been studied include MCI-Amnestic subtype, AD, bvFTD, aphasic stroke patients, and schizophrenia.[4-7]
Advantages	Easy to administer. It is well studied as a sub-test of D-KEFS, a highly standardized set of tests for many higher level functions.
Limitations	There is little normative data for proverb testing outside of larger testing batteries (like D-KEFS). It takes moderate amount of time and may be impractical at bedside. Results are commonly confounded by lack of familiarity with proverbs/idioms and/or impaired language function. It is an incomplete representation of executive dysfunction.
References	1. Gorham DR. Use of the proverbs test for differentiating schizophrenics from normals. J Consult Psychol 1956;20:435-440. 2. Barth A, Kufferle B. [Development of a proverb test for assessment of concrete thinking problems in schizophrenic patients]. Nervenarzt 2001;72:853-858. 3. Delis DC. Delis-Kaplan Executive Function System (D-KEFS). San Antonio, TX: The Psychological Corporation, 2001. 4. Leyhe T, Saur R, Eschweiler GW, Milian M. Impairment in proverb interpretation as an executive function deficit in patients with amnestic mild cognitive impairment and early Alzheimer's disease. Dement Geriatr Cogn Disord 2011;1:51-61. 5. Kempler D, Van Lancker D, Read S. Proverb and idiom comprehension in Alzheimer disease. Alzheimer Dis Assoc Disord 1988;2:38-49. 6. Van Lancker DR, Kempler D. Comprehension of familiar phrases by left- but not by right-hemisphere damaged patients. Brain Lang 1987;32:265-277. 7. Rapp AM, Wild B. Nonliteral language in Alzheimer dementia: A review. J Int Neuropsychol Soc 2011;17:207-218.

Test Name	Similarities Subtest, Wechsler Adult Intelligence Scale (WAIS)
Description	A verbal subtest of the Wechsler Adult Intelligence Scale (WAIS). It assesses abstract verbal reasoning, within the realms of verbal comprehension and executive function. The test is part of the widely used WAIS, which currently is in its 4th revision (WAIS-IV).[1] Testing similarities is a way of measuring some level of higher order conceptual thinking. It involves the ability to abstract meaning from a priori unrelated verbal information. Patients are asked, in a free response, how two words are alike. A set of items typically includes those with both simple and more abstract relationships. For example, a simple relationship may be that both items are a "parts of the body" (nose and tongue); a more complicated example could be that the two items both represent the "beginning stages of life" (bud and baby). Among other groups, patients with Alzheimer disease (AD) and schizophrenia have been studied extensively with similarities to test the executive function of verbal conceptual thinking.[2]
Specific Functions Assessed	Abstract verbal reasoning, verbal concept formation (directly); crystallized intelligence, auditory comprehension, associative and categorical thinking, verbal expression, and ability to distinguish between relevant and irrelevant features (indirectly)
Subscales	It itself is a subscale of the WAIS-IV; no subscales within
Number of Items/Scoring	It contains 18 items. Items are scored on a 0-2 point scale. A mean score range of about 18-28 for is expected for normals, across the lifespan. Raw scores are converted to standard scores, based upon the WAIS-IV scoring system, with a mean of 10 and a standard deviation of 3. Subtest short-forms of the WAIS-III are also available.[3]
Copyright Status	Copyrighted. Available through Pearson Assessments (The Psychological Corporation).
Administration Time	Approximately 3-5 minutes
Normative Psychometric Data	Normative data is available across the lifespan for the WAIS-IV and individual short-form Subtests of the WAIS-III, ages 6-90.[1, 3]
Sensitivity and Specificity	Vary across studies and is usually determined by the whole WAIS rather than the Similarities Subtest individually. Internal consistency reliability ranges have been 0.80 to 0.98 indicating a high level of internal consistency. It has been mostly studied in populations with MCI, AD, vascular dementia and schizophrenia.[2, 4, 5]
Advantages	Directions are simple to understand and it is easy to administer. It is a good measure for specific assessment of abstract verbal reasoning. It has

	robust normative psychometric data given that it is a subtest of a major testing battery.
Limitations	It may not be an accurate representation of patients' global executive function. Level of education may be a confounding factor in interpretation of results if the education-adjusted normative data is not used.
References	1. Wechsler D. Wechsler Adult Intelligence Scale, 4th ed. San Antonio, TX: Pearson, 2008. 2. Hays JR, Reas DL, Shaw JB. Concurrent validity of the Wechsler abbreviated scale of intelligence and the Kaufman brief intelligence test among psychiatric inpatients. Psychol Rep 2002;90:355-359. 3. Axelrod BN, Ryan JJ, Ward LC. Evaluation of seven-subtest short forms of the Wechsler Adult Intelligence Scale-III in a referred sample. Arch Clin Neuropsychol 2001;16:1-8. 4. Baudic S, Barba GD, Thibaudet MC, Smagghe A, Remy P, Traykov L. Executive function deficits in early Alzheimer's disease and their relations with episodic memory. Arch Clin Neuropsychol 2006;21:15-21. 5. Lamar M, Swenson R, Kaplan E, Libon DJ. Characterizing alterations in executive functioning across distinct subtypes of cortical and subcortical dementia. Clin Neuropsychol 2004;18:22-31.

Test Name	Boston Naming Test- 60 item
Description	A widely-used, visual confrontation naming test. It was first introduced by Kaplan, Goodglass, and Weintraub in 1983.[1] It was designed to detect word-retrieval difficulties in patients with neurologic disorders, like aphasia or Alzheimer disease. The test consists of 60 black and white drawings of various objects which are presented to the patient in increasing order of difficulty. Because there may be various appropriate answers, the patient is asked for the "common name" of each object. Patients may be asked for the common name again if there is no correct response within 20 seconds. Response latencies (e.g. 5, 10, and 15 seconds) are often recorded. Every response is recorded and often includes different types of errors (e.g. "wrong part" of object is given, "mispronunciation"). If there is no correct response after 20 seconds, a phonemic cue (e.g. "Dom-" for "Dominoes") may be offered. The cue may help differentiate deficits due to word-retrieval or search strategy from those due to a deficit in semantic knowledge.
Specific Functions Assessed	Visual processing, lexical-semantic memory and processing, output phonology, and speech.
Subscales	No subscales
Number of Items/Scoring	There are 60 items, with each item worth one point. Max score is 60 points. Multiple types of errors, including "wrong part" and "multiple attempts (where the correct response is *not* the last attempt) are coded. Patients are allowed 20 second to respond, after which they are given a semantic prompt, followed by a phonemic cue. If the correct name is given after the prompt only, this is indicated as +1 (for each name) to the score (e.g. 56/60, + 2 with prompts).[1]
Copyright Status	Copyrighted, Lippincott Williams & Wilkins; 2nd edition (2001)
Administration Time	10-20 minutes
Normative Psychometric Data	Available for all ages, with most comprehensive norms for adults, from ages 50-95.[2-7]
Sensitivity and Specificity	These vary across studies. Patient populations that have been studied the most include those with aphasia due to stroke, temporal lobectomy for epilepsy, MCI, AD, Huntington Disease, and Primary Progressive Aphasia (multiple variants).[8]
Advantages	It is easy to administer. It evaluates multiple language-related

	functions. It is a well-established and normed test.
Limitations	A strong cultural bias has been noted. It has a moderate length of administration time which may not always be practical for bedside assessment.
References	1. Kaplan E, Goodglass H, Weintraub S. Boston Naming Test. Philadelphia: Lea & Febiger, 1983. 2. Pedraza O, Graff-Radford NR, Smith GE, et al. Differential item functioning of the Boston Naming Test in cognitively normal African American and Caucasian older adults. J Int Neuropsychol Soc 2009;15:758-768. 3. Zec RF, Burkett NR, Markwell SJ, Larsen DL. Normative data stratified for age, education, and gender on the Boston Naming Test. Clin Neuropsychol 2007;21:617-637. 4. Steinberg BA, Bieliauskas LA, Smith GE, Langellotti C, Ivnik RJ. Mayo's older Americans normative studies: Age- and IQ-adjusted norms for the Boston Naming Test, the MAE Token Test, and the Judgment of Line Orientation Test. Clin Neuropsychol 2005;19:280-328. 5. Saxton J, Ratcliff G, Munro CA, et al. Normative data on the Boston Naming Test and two equivalent 30-item short forms. Clin Neuropsychol 2000;14:526-534. 6. Tombaugh TN, Hubley AM. The 60-item Boston Naming Test: Norms for cognitively intact adults aged 25 to 88 years. J Clin Exp Neuropsychol 1997;19:922-932. 7. Saykin AJ, Gur RC, Gur RE, et al. Normative neuropsychological test performance: Effects of age, education, gender and ethnicity. Appl Neuropsychol 1995;2:79-88. 8. Hodges JR, Salmon DP, Butters N. The nature of the naming deficit in Alzheimer's and Huntington's disease. Brain 1991;114 (Pt 4):1547-1558.

Test Name	Boston Naming Test (BNT- 15 item (Short Form)
Description	This test is a shorter variant of the BNT, a visual confrontation naming test designed to detect deficits in semantic retrieval and lexical access. Like the BNT-60 item version, the BNT-15 item Short Form assesses visual confrontation naming using black and white line drawings of common objects. There are 4 versions of the 15-item Short Form, which were created by assigning the 60 BNT items to four different item sets.[1] The items in each version are in a preserved order relative to the 60-item version. The 15 items in each set are also matched for word frequency, which itself is correlated to naming difficulty. Arranging the items beginning with easiest items first is felt to facilitate an accurate assessment of patients with limited attention spans. This form may be most useful for the serial assessments of semantic retrieval, and, for use in situations in which the administration of the complete BNT 60-item is not practical.
Specific Functions Assessed	Visual processing, lexical-semantic memory and processing, output phonology, and speech.
Subscales	No subscales
Number of Items/Scoring	It has a total of 15 items. Patients have 20 seconds to come up with the name for each item. Items are scored as correct, correct with a semantic cue, or correct with a phonemic cue. The total score is the number correct spontaneously, or, with semantic cues. A score of 13/15 was demonstrated as normal in a large cohort of healthy elderly controls (ages 65-93) with 12 or more years of education[2]. Education-adjusted scores for less than 12 years of education had about a 1 point decrease.
Copyright Status	Copyrighted, Psychological Assessment Resources, Inc. (PAR)
Administration Time	Typically 3-5 minutes
Normative Psychometric Data	Little normative data exist for the 15-item short versions. One cohort included 803 healthy adults, age 65-93.[2] Correlations with the full 60-item test range from .62 to .98, suggesting that the shortened versions are comparable to the 60-item test.[2]
Sensitivity and Specificity	Vary across studies and patient populations. Studies show differences in performance based on age and education, favoring younger and more educated people. Although the effects of age and education on performance are not clear, education has been shown to have a

	greater impact than age.[3] It has mostly been studied in patients with AD[4] and AD/Vascular dementia.[5]
Advantages	It is a brief, simple screen of patient's naming performance and capacity to retrieve semantic knowledge. Spanish versions are available.[1]
Limitations	It does not fully assess a patient's language abilities and deficits. Because language or semantic impairment is common in early AD, it may not be an appropriate screening tool for a level of AD dementia. Demographic factors, including language, education, and culture, have been shown to affect one's performance. Longitudinal data are limited.
References	1. Mack WJ, Freed DM, Williams BW, Henderson VW. Boston Naming Test: Shortened versions for use in Alzheimer's disease. J Gerontol 1992;47:P154-158. 2. Kent PS, Luszcz MA. A review of the Boston Naming Test and multiple-occasion normative data for older adults on 15-item versions. Clin Neuropsychol 2002;16:555-574. 3. Mitrushina M. Handbook of normative data for neuropsychological assessment: Oxford University Press, USA, 2005. 4. Lansing AE, Ivnik RJ, Cullum CM, Randolph C. An empirically derived short form of the Boston naming test. Arch Clin Neuropsychol 1999;14:481-487. 5. Graves RE, Bezeau SC, Fogarty J, Blair R. Boston Naming Test short forms: A comparison of previous forms with new item response theory based forms. J Clin Exp Neuropsychol 2004;26:891-902.

Test Name	Controlled Oral Word Association Test (COWAT)
Description	The COWAT is a test of verbal phonemic fluency and generative capacity. Designed in the background of earlier verbal fluency tasks since the 1930s (like the Chicago Word Fluency test),[1] the COWAT has been widely used since its introduction in the 1980s[2] to assess verbal communication abilities in both healthy adults and those with various brain disorders. It has also been used to monitor delay of language development in children. Patients are given 1 minute to generate as many words as possible with a starting letter (e.g., 'F', or 'A'). The same 1 minute protocol is carried out for several (often 3) phonetic subcategories (i.e., letters). A common version, FAS, includes a total of 3 trials, with the letters 'F', 'A', and 'S'. It is thought to assess aspects of both language and executive or frontal-lobe function.
Specific Functions Assessed	Search of lexicon based on the first letter of words, output phonology, generative capacity
Subscales	It exists as a sub-test of the Multilingual Aphasia Examination (MAE).[3]
Number of Items/Scoring	The most common performance measure is the total number of unique words produced for all three letters. Other measures, such as number of words produced for each letter, or error patterns such as intrusions (words from the previous letter) and word repetitions (reflecting a failure in self-monitoring) have also been used.
Copyright Status	Copyrighted, Psychological Assessment Resources, Inc. (PAR)
Administration Time	5 minutes
Normative Psychometric Data	Available for all ages, with most comprehensive norms for adults, ages 20 to 90.[4-7]
Sensitivity and Specificity	Vary across studies. It has been studied in patients with aphasia, MCI and AD dementia,[8, 9] and patients following surgery for frontal lobe epilepsy.[10] It has also been studied as a measure of language and executive function in schizophrenia[11] and the frontotemporal dementias.[12]
Advantages	It is easy to administer, brief, and may be repeated with low concern for bias or practice effects. It has robust normative data across the lifespan.
Limitations	It is important to use age- and education -adjusted norms, as results can be impacted by level of education and literacy. As a language

	test, it is significantly affected by executive dysfunction. As a test of executive function, can be significantly affected by language dysfunction.
References	1. Thurstone LL, Thurstone TG. Primary mental abilities. Chicago: University of Chicago Press, 1938. 2. Benton AL, Hamsher K. Multilingual Aphasia Examination. Iowa City: AJA Associates, 1989. 3. Benton AL, A.B. S, Hamsher K, Varney NR, Spreen O. Contributions to neuropsychological assessment: A clinical manual., 2nd ed. New York: Oxford University Press, 1994. 4. Steinberg BA, Bieliauskas LA, Smith GE, Ivnik RJ. Mayo's older Americans normative studies: Age- and IQ-adjusted norms for the Trail-Making Test, the Stroop Test, and MAE Controlled Oral Word Association Test. Clin Neuropsychol 2005;19:329-377. 5. Loonstra AS, Tarlow AR, Sellers AH. COWAT metanorms across age, education, and gender. Appl Neuropsychol 2001;8:161-166. 6. Ross TP, Calhoun E, Cox T, Wenner C, Kono W, Pleasant M. The reliability and validity of qualitative scores for the Controlled Oral Word Association Test. Arch Clin Neuropsychol 2007;22:475-488. 7. Ruff RM, Light RH, Parker SB, Levin HS. Benton Controlled Oral Word Association Test: Reliability and updated norms. Arch Clin Neuropsychol 1996;11:329-338. 8. Rinehardt E, Eichstaedt K, Schinka JA, et al. Verbal fluency patterns in mild cognitive impairment and Alzheimer's disease. Dement Geriatr Cogn Disord 2014;38:1-9. 9. Henry JD, Crawford JR, Phillips LH. Verbal fluency performance in dementia of the Alzheimer's type: A meta-analysis. Neuropsychologia 2004;42:1212-1222. 10. Sarkis RA, Busch RM, Floden D, et al. Predictors of decline in verbal fluency after frontal lobe epilepsy surgery. Epilepsy Behav 2013;27:326-329. 11. Kremen WS, Seidman LJ, Faraone SV, Tsuang MT. Is there disproportionate impairment in semantic or phonemic fluency in schizophrenia? J Int Neuropsychol Soc 2003;9:79-88. 12. Roca M, Manes F, Gleichgerrcht E, et al. Intelligence and executive functions in frontotemporal dementia. Neuropsychologia 2013;51:725-730.

Test Name	Semantic/Category Fluency Test
Description	A test of verbal semantic fluency and generative language capacity. It is probably the most widely used measure of semantic/category verbal fluency. It often given along with the controlled oral word association test (COWAT),[1] which is thought to assess the "phonemic" rather than "semantic" aspect of verbal fluency. In this test, used broadly since the 1970s, patients are given 1 minute to provide as many examples that they are able to within a certain category (e.g., "animals", "grocery store items", "vegetables").[2] It is sometimes referred to as a "free-listing" task because patients are not bound by any phonetic cue or guideline for word-generation. Their "search strategy" to retrieve words is based on the meaning of words and the clustering of words with a similar meaning, rather than by the way a word sounds (phonology) when produced.
Specific Functions Assessed	Lexical-semantic processing, organization and strategy of thought, output phonology, aspects of speech
Subscales	No subscales
Number of Items/Scoring	The most common performance measure is the total number of words across all the categories used. Typical administration involves trials with 3 different categories (e.g. "animals," "vegetables," "fruits"). The score is often the aggregate of total words over the 3 categories. Other analyses, such as number of repetitions or number of switches to different clusters or sub-categories of words, can be carried out.
Copyright Status	Public domain
Administration Time	5 minutes
Normative Psychometric Data	Available for all ages with most comprehensive norms for adults.[3-6]
Sensitivity and Specificity	Vary across studies. Patients with AD, MCI, and schizophrenia may be the most commonly studied with this instrument.[7-9]
Advantages	It is easy to brief and easy to administer. It can be repeated with low concern for bias or practice effects. It is well studied and has ample normative data.
Limitations	It can be biased by the educational background of participants. Performance may be confounded by executive function deficits.

References	1. Benton AL, Hamsher K. Multilingual Aphasia Examination. Iowa City: AJA Associates, 1989. 2. Ardila A, Ostrosky-Solís F, Bernal B. Cognitive testing toward the future: The example of semantic verbal fluency (animals). International Journal of Psychology 2006;41:324-332. 3. Holtzer R, Goldin Y, Zimmerman M, Katz M, Buschke H, Lipton RB. Robust norms for selected neuropsychological tests in older adults. Arch Clin Neuropsychol 2008;23:531-541. 4. Acevedo A, Loewenstein DA, Barker WW, et al. Category fluency test: Normative data for English- and Spanish-speaking elderly. J Int Neuropsychol Soc 2000;6:760-769. 5. Gladsjo JA, Schuman CC, Evans JD, Peavy GM, Miller SW, Heaton RK. Norms for letter and category fluency: Demographic corrections for age, education, and ethnicity. Assessment 1999;6:147-178. 6. 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;19:243-269. 7. Sumiyoshi C, Matsui M, Sumiyoshi T, Yamashita I, Sumiyoshi S, Kurachi M. Semantic structure in schizophrenia as assessed by the category fluency test: Effect of verbal intelligence and age of onset. Psychiatry Res 2001;105:187-199. 8. Teng E, Leone-Friedman J, Lee GJ, et al. Similar verbal fluency patterns in amnestic mild cognitive impairment and Alzheimer's disease. Arch Clin Neuropsychol 2013;28:400-410. 9. Clark LJ, Gatz M, Zheng L, Chen YL, McCleary C, Mack WJ. Longitudinal verbal fluency in normal aging, preclinical, and prevalent Alzheimer's disease. Am J Alzheimers Dis Other Demen 2009;24:461-468.

Test Name	Western Aphasia Battery-Revised
Description	A battery of tests that assess language performance. It was designed as a broad detection tool to look for deficits in the major clinical aspects of language function (e.g., auditory comprehension, reading, writing).[1] The revised version (WAB-R, 2006) includes additional tests that discriminate between different types of dyslexia. The battery is designed to provide a broad assessment of both language ability and intellectual ability. The Aphasia Quotient (AQ) is one of the generated scores which can be thought of as an aggregate measure of language function. The Cortical Quotient (CQ) is viewed as a more general measure of intellectual ability as it reflects results from all 8 subtests. The CQ takes into account drawing, block design, praxis, and calculation as well as language.[2]
Specific Functions Assessed	Many aspects of language function, including: content, fluency, auditory comprehension, repetition, naming, reading, and writing; also assesses praxis and calculation
Subscales	8 subtests (32 short tasks)
Number of Items/Scoring	Aphasia Quotient; Cortical Quotient; Auditory Comprehension Quotient; Oral Expression Quotient; Reading Quotient; Writing Quotient
Copyright Status	Copyrighted, Pearson PLC
Administration Time	45-60 minutes
Normative Psychometric Data	Available for all ages (ages 18-89), with the most comprehensive norms for adults.[3, 4]
Sensitivity and Specificity	Both sensitive and specific for aphasic deficits if entire battery is administered. It has been studied in populations with aphasia as a result of stroke, head injury, or neurodegenerative disease, including those with Alzheimer disease and primary progressive aphasia (PPA).[5, 6]
Advantages	Multiple language functions are evaluated, which provides a broad-based language assessment. It has been well-studied and has ample normative data.
Limitations	Lengthy administration time and is thus often impractical for bedside testing. Often specific subtests only are used due to limited time for testing. Normative values and sensitivity and specificity are only relevant for the entire battery.

References	1. Kertesz A, McCabe P. Recovery patterns and prognosis in aphasia. Brain 1977;100 Pt 1:1-18. 2. Spreen O, Risser A. Assessment of aphasia. New York: Oxford University Press, 2003. 3. Shewan CM, Kertesz A. Reliability and validity characteristics of the Western Aphasia Battery (WAB). J Speech Hear Disord 1980;45:308-324. 4. Kim H, Na DL. Normative data on the Korean version of the Western Aphasia Battery. J Clin Exp Neuropsychol 2004;26:1011-1020. 5. Code C, Rowley D, Kertesz A. Predicting recovery from aphasia with connectionist networks: Preliminary comparisons with multiple regression. Cortex 1994;30:527-532. 6. Blair M, Marczinski CA, Davis-Faroque N, Kertesz A. A longitudinal study of language decline in Alzheimer's disease and frontotemporal dementia. J Int Neuropsychol Soc 2007;13:237-245.

Memory

Test Name	California Verbal Learning Test (CVLT)- II
Description	A word-list verbal learning and episodic memory test. It was first introduced in 1987 (now called CVLT-I) and was subsequently shown to be sensitive to memory deficits in a broad range of conditions.[1] It includes a word-list learning task that examines aspects of memory registration, recall strategy, and memory storage. The CVLT II, a more recent, updated version (2000), involves the same task, but includes the addition of a forced choice trial to assess level of effort, and, the inclusion of recall discriminability indices, which take into account intrusion errors.[2] It has had especially widespread use in the evaluation of learning and memory following traumatic brain injury. In the test, patients learn 16 words over 5 learning trials, and then have free and cued recall after a delayed period. Recognition memory is then tested with a forced-choice, "yes-no" recognition task.[3]
Specific Functions Assessed	Assesses multiple cognitive components of verbal learning and memory functioning, including recall strategy/organization, interference effects, cuing and recognition, and recall accuracy and effort
Subscales	No subscales
Number of Items/Scoring	The *standard form* has 16 words representing 4 categories. It includes 5 learning trials, with immediate and long-delayed free and cued recall and "yes-no" recognition. The *alternate form* includes an immediate recall trial, using a second "interference" word-list, which is given in-between the immediate and delayed recall trials of the original list. A s*hort form* has 9 words representing 3 categories. It includes learning trials with immediate recall, followed by a counting backwards task to provide interference, and then delayed recall with a shorter interval. Comprehensive data analysis is achieved via computerized scoring.
Copyright Status	Copyrighted. Available through Pearson PLC
Administration Time	*Standard and alternate Forms:* 30 minutes testing, plus 30 minutes for delay; *Short form:* 15 minutes testing, plus 15 minutes for delay
Normative Psychometric Data	The CVLT-II underwent national standardization through the study of 1,087 individuals (565 females), well matched to the most recent U.S. Census in race/ethnicity, education level, and geographic region. Norms were provided for seven age groups, ages 16 to 89.[1] Separate norms are provided for males and females because of gender differences in total recall. Test-retest and practice effects were evaluated on standard and alternate forms using a normative sample.[4]

	Additional normative values are available across the lifespan.[5]
Sensitivity and Specificity	In a study comparing different memory tests, the CVLT-II ranked highest in terms of distinguishing MCI from normal aging (sensitivity = 90.2; specificity = 84.2).[6] The test has been studied most in populations of MCI and AD,[6-8] but also in those with psychiatric disorders, multiple sclerosis, and traumatic brain injury.[9, 10]
Advantages	It provides detailed indices regarding learning strategies and organization. It has been used in a large number of studies for evaluating learning and memory function in a wide array of neurologic and psychiatric disorders, particularly traumatic brain injury and dementia. The short (9-word) form may be particularly useful as a screening evaluation for dementia patients, but may not be as sensitive to mild deficits as the standard form.
Limitations	It is copyrighted and interpretation does involve computerized scoring algorithms. The CVLT-II short from (9-word list) has less specificity as a memory test than the standard 16-word list.
References	1. Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test. San Antonio, TX: The Psychological Corporation, 1987. 2. Delis DC, Kramer JH, Kaplan E, Ober BA. California Verbal Learning Test–II, 2nd ed. San Antonio, TX: The Psychological Corporation, 2000. 3. Elwood RW. The California Verbal Learning Test: Psychometric characteristics and clinical application. Neuropsychol Rev 1995;5:173-201. 4. Woods SP, Delis DC, Scott JC, Kramer JH, Holdnack JA. The California Verbal Learning Test--second edition: Test-retest reliability, practice effects, and reliable change indices for the standard and alternate forms. Arch Clin Neuropsychol 2006;21:413-420. 5. Fine EM, Kramer JH, Lui LY, Yaffe K, Study Of Osteoporotic Fractures Sof Research Group. Normative data in women aged 85 and older: Verbal fluency, digit span, and the CVLT-II short form. Clin Neuropsychol 2012;26:18-30. 6. Rabin LA, Pare N, Saykin AJ, et al. Differential memory test sensitivity for diagnosing amnestic mild cognitive impairment and predicting conversion to Alzheimer's disease. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2009;16:357-376. 7. Beck IR, Gagneux-Zurbriggen A, Berres M, Taylor KI, Monsch AU. Comparison of verbal episodic memory measures: Consortium to Establish a Registry for Alzheimer's Disease--Neuropsychological Assessment Battery (CERAD-NAB) versus California Verbal Learning Test (CVLT). Arch Clin Neuropsychol 2012;27:510-519.

	8. Lacritz LH, Cullum CM. The Hopkins Verbal Learning Test and CVLT: A preliminary comparison. Arch Clin Neuropsychol 1998;13:623-628. 9. Stegen S, Stepanov I, Cookfair D, et al. Validity of the California Verbal Learning Test-II in multiple sclerosis. Clin Neuropsychol 2010;24:189-202. 10. Arenth PM, Russell KC, Scanlon JM, Kessler LJ, Ricker JH. Encoding and recognition after traumatic brain injury: Neuropsychological and functional magnetic resonance imaging findings. J Clin Exp Neuropsychol 2012;34:333-344.

Test Name	CERAD-Word List Memory Test
Description	A measure of verbal learning and memory. It is part of a larger neuropsychological battery, created for the Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Other CERAD sub-tests include the Boston Naming Test (BNT) and a constructional praxis test, among a total of 8 subtests.[1] The CERAD-Word List Memory test has been widely used since the 1980s. It employs a word-list learning task to examine aspects of verbal learning and memory. The patient has 3 learning trials to learn 10 words, during which patients sequentially read one word aloud every 2 seconds. Patients are then asked for both delayed recall and forced-choice recognition after a time-delay.
Specific Functions Assessed	Verbal episodic memory, auditory working/episodic memory function
Subscales	No subscales
Number of Items/Scoring	It includes a 10-word list of unrelated items, with 3 learning trials, a delayed recall trial, and a forced-choice recognition recall trial, which includes 10 targets mixed in with 10 distracters. The scores are: the number of learned words (out of 30) in the learning trials, the number of freely recalled words (out of 10) in the delayed trial, and, the number of correct targets (minus false positives) identified in the forced-choice trial.
Copyright Status	Public domain, but materials cost a fee; the CERAD website has detailed information and references (http://cerad.mc.duke.edu)
Administration Time	5-10 minutes total, with 5-8 minute interval between last learning trial and delayed recall
Normative Psychometric Data	Normative data were obtained with 958 carefully screened, nationally distributed white and black patients with AD and 413 non-demented controls.[2, 3] Validation studies demonstrate good-test-retest reliability, discriminant validity, and longitudinal stability (normal elderly controls).[4, 5] Large population studies have been conducted in Spanish, German, Chinese, Korean, Arabic, Finnish, Japanese, and other populations.
Sensitivity and Specificity	Delayed recall may be the best way to discriminate between normal elderly controls (94% accurate) and patients with mild AD (86% accurate). The test does not reliably distinguish AD severity levels (floor effect). It has been most studied in the AD population, but has also been studied in MCI and PD.[6-8]
Advantages	It is well-validated and easy to administer in the clinical context. It is available in multiple languages. Abundant data are available as part

	of comprehensive diagnostic assessment and longitudinal study of AD.
Limitations	It may not be as sensitive in detecting MCI as other word-list tests. There are limited data in non-AD populations.
References	1. Morris JC, Heyman A, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part I. Clinical and neuropsychological assessment of Alzheimer's disease. Neurology 1989;39:1159-1165. 2. Welsh KA, Butters N, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. Neurology 1994;44:609-614. 3. Welsh KA, Fillenbaum G, Wilkinson W, et al. Neuropsychological test performance in African-American and white patients with Alzheimer's disease. Neurology 1995;45:2207-2211. 4. Fillenbaum GG, Unverzagt FW, Ganguli M, Welsh-Bohmer KA, Heyman A. The CERAD neuropsychological battery: Performance of representative community and tertiary care samples of African American and European American elderly. In: Ferraro FR, ed. Minority and cross cultural aspects of neuropsychological assessment: Taylor & Francis, 2002. 5. Stein J, Luppa M, Maier W, et al. The assessment of changes in cognitive functioning in the elderly: Age- and education-specific reliable change indices for the SIDAM. Dement Geriatr Cogn Disord 2012;33:73-83. 6. Beck IR, Gagneux-Zurbriggen A, Berres M, Taylor KI, Monsch AU. Comparison of verbal episodic memory measures: Consortium to Establish a Registry for Alzheimer's Disease--Neuropsychological Assessment Battery (CERAD-NAB) versus California Verbal Learning Test (CVLT). Arch Clin Neuropsychol 2012;27:510-519. 7. Welsh K, Butters N, Hughes J, Mohs R, Heyman A. Detection of abnormal memory decline in mild cases of Alzheimer's disease using CERAD neuropsychological measures. Arch Neurol 1991;48:278-281. 8. Welsh KA, Butters N, Hughes JP, Mohs RC, Heyman A. Detection and staging of dementia in Alzheimer's disease. Use of the neuropsychological measures developed for the Consortium to Establish a Registry for Alzheimer's Disease. Arch Neurol 1992;49:448-452.

Test Name	Hopkins Verbal Learning Test-Revised (HVLT-R)
Description	A verbal learning and episodic memory test. The original HVLT was developed in 1991[1] as a repeatable, screening measure of verbal learning and memory. It was designed as a brief test, useful for clinical situations where a lengthy comprehensive assessment is not practical. This may include testing patients with AD or other chronic memory disorders. The original HVLT included a list of 12 words, from 3 different categories, where patients engage in 3 learning trials and then an immediate free recall trial. The revised version, HVLT-R,[2] retained the original words and procedure and added a delayed verbal recall and recognition trial.
Specific Functions Assessed	Verbal learning and episodic memory
Subscales	No subscales
Number of Items/Scoring	There are many forms. Each form consists of a list of 12 nouns (targets) with four words drawn from each of three semantic categories. The semantic categories differ across the 6 forms, but the forms are very similar in their psychometric properties. Six alternate forms are available to minimize practice effects with repeated administration. Among the 12 foils in the recognition set, 6 are semantically-related to target words (2 in each of 3 categories) and the other 6 are unrelated. False positive errors are more likely to occur with semantically-related foils in normal control patients. Raw scores are derived for Total Recall, Delayed Recall, Retention (percent retained), and a Recognition Discrimination Index.
Copyright Status	Copyrighted, Psychological Assessment Resources, Inc. (PAR)
Administration Time	5-10 minutes for 3 learning trials, delayed recall, and forced-choice recognition, with a 20-25 minute delay between last learning trial and delayed recall
Normative Psychometric Data	Ample normative data exist for ages 17 to late 80s.[2] It has high test-retest reliability, and it has well established construct, concurrent, and discriminant validity.[2-5]
Sensitivity and Specificity	HVLT total learning score exhibits sensitivity and specificity of 87% and 98%, respectively, for discriminating patients with dementia from healthy controls, with an optimal discriminative capacity between patients with MCI and those with normal cognition (NC).[6-8] It has also been studied in TBI and vascular dementia patients.
Advantages	It is easy to administer and score and is well-tolerated, even by significantly impaired individuals. It has been validated within brain-disordered populations (e.g., Alzheimer's disease, amnestic disorders).

	Recent validation studies using Spanish[9] and Chinese[10] versions show comparable ability to distinguish dementia from healthy elderly controls.
Limitations	It is copyrighted. It may not be as sensitive for MCI compared to other, more complex word-list tests and it may have ceiling effects in younger patients.
References	1. Brandt J. The Hopkins Verbal Learning Test: Development of a new memory test with six equivalent forms. Clin Neuropsychol 1991;5:125-142. 2. Benedict RH, Schretlen D, Groninger L, Brandt J. The Hopkins Verbal Learning Test-Revised: Normative data and analysis of interform and test-retest reliability. Clin Neuropsychol 1998;12:43-55. 3. Lacritz LH, Cullum CM. The Hopkins Verbal Learning Test and CVLT: A preliminary comparison. Arch Clin Neuropsychol 1998;13:623-628. 4. Rasmusson DX, Bylsma FW, Brandt J. Stability of performance on the Hopkins Verbal Learning Test. Arch Clin Neuropsychol 1995;10:21-26. 5. Shapiro AM, Benedict RH, Schretlen D, Brandt J. Construct and concurrent validity of the Hopkins Verbal Learning Test-Revised. Clin Neuropsychol 1999;13:348-358. 6. de Jager CA, Schrijnemaekers AC, Honey TE, Budge MM. Detection of MCI in the clinic: Evaluation of the sensitivity and specificity of a computerised test battery, the Hopkins Verbal Learning Test and the MMSE. Age Ageing 2009;38:455-460. 7. Frank RM, Byrne GJ. The clinical utility of the Hopkins Verbal Learning Test as a screening test for mild dementia. Int J Geriatr Psychiatry 2000;15:317-324. 8. Hogervorst E, Combrinck M, Lapuerta P, Rue J, Swales K, Budge M. The Hopkins Verbal Learning Test and screening for dementia. Dement Geriatr Cogn Disord 2002;13:13-20. 9. Gonzalez-Palau F, Franco M, Jimenez F, Parra E, Bernate M, Solis A. Clinical utility of the Hopkins Verbal Test-Revised for detecting Alzheimer's disease and mild cognitive impairment in Spanish population. Arch Clin Neuropsychol 2013;28:245-253. 10. 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.

Test Name	Logical Memory Test (I and II)
Description	A test of episodic verbal memory function. It is a subtest of the Wechsler Memory Scale (WMS)-IV[1]. This current format, WMS-IV, is an adapted version of prior scales, the first of which appeared in 1945 as the original Wechsler Memory Scale (WMS). This subtest is also known as the Immediate and Delayed Paragraph Recall test and other paragraph recall tests with a similar format have been studied. It has been used widely to study patient populations with dementia and TBI among other neurologic disorders. The test has 2 parts. In part I, the patient is asked to freely recall as many details from 2 short narratives as possible, immediately after each paragraph is read aloud. In part II, the patient is asked to freely recall these details after a 20- 30 minute delay. After the delayed free recall, there is a recognition-memory task, where the patient is asked 15 "yes/no" questions about the content of each paragraph.
Specific Functions Assessed	Immediate and delayed auditory verbal episodic memory (directly); language comprehension and auditory working memory (indirectly)
Subscales	No subscales; it is a subtest of a larger standardized memory scale (WMS-IV)
Number of Items/Scoring	It has an Adult version (age 16-69) and Older Adult version (age 65-90). On the Older Adult, the first narrative is shorter than the second one, where the narratives are equal length on the Adult version. The range of raw scores from the Adult Version of Logical Memory I and II Recall is 0 to 50 points each. Adult version Logical Memory Recognition total raw scores range from 0 to 30 points. Raw scores may be converted to scaled scores using the normative tables in the WMS-IV Scoring Manual. Scaled scores range from 1 to 19.
Copyright Status	Copyrighted, Pearson PLC
Administration Time	10-15 minutes total, with 20-30 minute delay between parts I and II
Normative Psychometric Data	There is ample normative data across the lifespan, included in the WMS-IV.[2] These Data load heavily on same factor as Verbal Paired Associates (another WMS-IV subtest) in a factor analysis. The overall internal-consistency reliability coefficients of the normative samples are in the 0.80s and 0.90s.
Sensitivity and Specificity	The Logical Memory Test cutoffs (adjusted for education) used by ADNI and similar studies (e.g., -1.5 standard deviations for late MCI) are widely used. The Logical Memory Test scales with AD-8 in Alzheimer disease.

Advantages	It is easily administered and scored, but does require training. Multiple versions are available. Logical Memory Story recall is widely used in clinical trials for MCI and AD dementia. It use alone or in conjunction with other tests, such as word-list learning tasks (e.g. California Verbal Learning Test), has shown utility in predicting MCI to AD conversion,[3, 4] response to treatment in depression[5] and outcomes in Traumatic brain Injury.[6] It may also predict driving performance,[7] and treatment decisional abilities in mild to moderate dementia[8] and correlates with outcomes in many diverse neurological, medical and psychiatric conditions.[9]
Limitations	It is copyrighted. Practice effects can interfere with longitudinal assessment. Normative scores do vary with educational attainment
References	1. Wechsler D. Wechsler Adult Intelligence Scale, 4th ed. San Antonio, TX: Pearson, 2008. 2. Wechsler D. Wechsler Memory Scale 4th ed. San Antonio, TX: Pearson, 2008. 3. Rabin LA, Pare N, Saykin AJ, et al. Differential memory test sensitivity for diagnosing amnestic mild cognitive impairment and predicting conversion to Alzheimer's disease. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 2009;16:357-376. 4. Tatsuoka C, Tseng H, Jaeger J, et al. Modeling the heterogeneity in risk of progression to Alzheimer's disease across cognitive profiles in mild cognitive impairment. Alzheimers Res Ther 2013;5:14. 5. Story TJ, Potter GG, Attix DK, Welsh-Bohmer KA, Steffens DC. Neurocognitive correlates of response to treatment in late-life depression. Am J Geriatr Psychiatry 2008;16:752-759. 6. Green RE, Colella B, Hebert DA, et al. Prediction of return to productivity after severe traumatic brain injury: Investigations of optimal neuropsychological tests and timing of assessment. Arch Phys Med Rehabil 2008;89:S51-60. 7. Szlyk JP, Myers L, Zhang Y, Wetzel L, Shapiro R. Development and assessment of a neuropsychological battery to aid in predicting driving performance. J Rehabil Res Dev 2002;39:483-496. 8. Gurrera RJ, Moye J, Karel MJ, Azar AR, Armesto JC. Cognitive performance predicts treatment decisional abilities in mild to moderate dementia. Neurology 2006;66:1367-1372. 9. Xiang YT, Shum D, Chiu HF, Tang WK, Ungvari GS. Association of demographic characteristics, symptomatology, retrospective and prospective memory, executive functioning and intelligence with social functioning in schizophrenia. Aust N Z J Psychiatry 2010;44:1112-1117.

Test Name	NYU Paragraph Recall Test
Description	A test of auditory, episodic verbal memory. It involves the immediate and delayed recall of a brief paragraph that is read aloud to a patient. It is the paragraph sub-test of a memory assessment called the Guild Memory Test,[1] first introduced in 1968. The Guild Memory Test, designed to differentiate normal cognitive aging from neurologic disease in the elderly, focuses on contextual memory and paired or associative learning function. The NYU Paragraph Recall Test has been used extensively by itself in both clinical research and drug therapy trials for MCI and AD.[2, 3]
Specific Functions Assessed	Episodic verbal memory (directly); language comprehension (indirectly)
Subscales	No subscales; it is itself a subtest of the Guild Memory Test.
Number of Items/Scoring	20 items
Copyright Status	Copyrighted (as a subtest of the Guild Memory Test)
Administration Time	A total of 12-15 minutes, with 5 minutes for Immediate recall and 2 minutes for delayed recall, with at least a 5 minute interval between
Normative Psychometric Data	Some normative data exists for the immediate and delayed recall trials, separately. One study of age-related cognitive decline includes 369 normal adults, age range 42-90 (mean age 62).[4] These data include results for normal controls compared to MCI and AD; values are not sub-divided by age.
Sensitivity and Specificity	One study demonstrates a 90% sensitivity in distinguishing healthy older adults from the cognitively impaired (using -1 SD from mean).[2] Other results, using a regression analysis and a cutoff score of 6, predicted decline from non-demented controls to AD dementia with 78-93% specificity and 82% sensitivity. The same cut-off score had 83% specificity and 96% sensitivity, with an accuracy of 92%, in predicting decline from MCI to AD.
Advantages	It is a brief, sensitive test, especially for older age groups. It is similar to the Logical Memory Test (from WMS-IV), which has been well studied and validated. Some normative data exists for immediate recall only, which can make administration time shorter. English and Spanish versions have been developed.
Limitations	It has not been used in routine clinical practice. Extensive normative data does not exist Does not reliably differentiate depression from neurological disorders.

References	1. Catalano FL. A preliminary report on a new memory scale. Percept Mot Skills 1968;27:277-278. 2. Kluger A, Ferris SH, Golomb J, Mittelman MS, Reisberg B. Neuropsychological prediction of decline to dementia in nondemented elderly. J Geriatr Psychiatry Neurol 1999;12:168-179. 3. Convit A, De Leon MJ, Tarshish C, et al. Specific hippocampal volume reductions in individuals at risk for Alzheimer's disease. Neurobiol Aging 1997;18:131-138. 4. De Santi S, Pirraglia E, Barr W, et al. Robust and conventional neuropsychological norms: Diagnosis and prediction of age-related cognitive decline. Neuropsychology 2008;22:469-484.

Test Name	Rey Auditory Verbal Learning Test (RAVLT)
Description	An auditory episodic verbal memory test that involves a word-list learning paradigm. It was originally created in the 1960s by Andre Rey as an adaptation of many prior word-list learning tests.[1] Word-list learning is argued to be among the most sensitive verbal memory test formats because the patient is freed from distraction by any associative context (e.g., a narrative structure). The RAVLT has been studied extensively in healthy cognitive aging groups as well as those who have memory impairment along the spectrum of AD.
Specific Functions Assessed	Verbal learning and episodic memory (directly); rote memorization, proactive and retroactive interference, retention, encoding versus retrieval, and subjective organization (indirectly)
Subscales	No subscales
Number of Items/Scoring	A 15-item word list (A) is presented, with 5 learning trials. Responses are recorded in the order provided by the patient. A summary score is calculated for the total number of words recalled across the 5 trials. Then, a second 15-word list (B) is presented to the patient, followed by a 6th presentation of the initial word list (A). If immediate recall for A on this trial is less than 13, then a 50-word recognition list is given in a forced-choice manner. These 50 words include all 30 words on lists A and B, and words that are either semantically related, phonologically related, or both, to a subset of words on Lists A and B. Alternatively, delayed recall of List A at 30 minutes or more can be administered to assess retention.
Copyright Status	Public Domain
Administration Time	15 minutes, with a 30 minute period before delayed recall
Normative Psychometric Data	Normative data has been acquired using multiple forms in patients age 8 to older than 80 in a number of different countries and in different clinical populations.[2-4] Test-retest reliability and criterion validity have generally been good.[5] An abridged form of the RAVLT is a component of the NIH Toolbox, which includes the summed recall of a 15-word list administered over 3 learning trials (NIHtoolbox.org).
Sensitivity and Specificity	It is sensitive in distinguishing a variety of developmental (e.g. ADHD, learning disabilities) and acquired brain disorders (e.g. MCI, AD, TBI)[6-10] from age-adjusted healthy controls. Delayed recall has also been shown to be a robust predictor of dementia in a large community sample followed for 5 and 10 years.[11]
Advantages	It has more extensive word-list length and number of learning trials than other tests (CERAD, Hopkins-VLT). It has relatively large amount of validation data. Offers some flexibility in the depth and degree of

	memory assessment. Included as part of the Alzheimer Disease Neuroimaging Initiative (ADNI) neuropsychological test battery, which offers a large dataset in conjunction with brain imaging and CSF biomarkers. Has been shown to be most sensitive ADNI cognitive test to early AD deficits.
Limitations	It may be more suitable for research applications or formal psychometric assessment than a bedside clinical tool.
References	1. Rey A. L'examen psychologique dans les cas d'encephalopathie traumatique. Arch Psychol (Geneve) 1941;28:286-340. 2. Ivnik RJ, Malec JF, Smith GE, et al. Mayo's older Americans normative studies: Updated AVLT norms for ages 56 to 97. The Clinical Neuropsychologist 2002;6 (Supplement):83-104. 3. Shapiro DM, Harrison DW. Alternate forms of the AVLT: A procedure and test of form equivalency. Arch Clin Neuropsychol 1990;5:405-410. 4. Uchiyama CL, D'Elia LF, Dellinger AM, et al. Alternate forms of the Auditory-Verbal Learning Test: Issues of test comparability, longitudinal reliability, and moderating demographic variables. Arch Clin Neuropsychol 1995;10:133-145. 5. Geffen GM, Butterworth P, Geffen LB. Test-retest reliability of a new form of the auditory verbal learning test (AVLT). Arch Clin Neuropsychol 1994;9:303-316. 6. Binder LM, Villanueva MR, Howieson D, Moore RT. The Rey AVLT recognition memory task measures motivational impairment after mild head trauma. Arch Clin Neuropsychol 1993;8:137-147. 7. Estevez-Gonzalez A, Kulisevsky J, Boltes A, Otermin P, Garcia-Sanchez C. Rey verbal learning test is a useful tool for differential diagnosis in the preclinical phase of Alzheimer's disease: Comparison with mild cognitive impairment and normal aging. Int J Geriatr Psychiatry 2003;18:1021-1028. 8. Pollak Y, Kahana-Vax G, Hoofien D. Retrieval processes in adults with ADHD: A RAVLT study. Dev Neuropsychol 2008;33:62-73. 9. Ricci M, Graef S, Blundo C, Miller LA. Using the Rey auditory verbal learning Test (RAVLT) to differentiate Alzheimer's dementia and behavioural variant fronto-temporal dementia. Clin Neuropsychol 2012;26:926-941. 10. Talley JL. Memory in learning disabled children: Digit span and the Rey auditory verbal learning Test. Arch Clin Neuropsychol 1986;1:315-322. 11. Tierney MC, Moineddin R, McDowell I. Prediction of all-cause dementia using neuropsychological tests within 10 and 5 years of diagnosis in a community-based sample. J Alzheimers Dis

	2010;22:1231-1240.

Test Name	Verbal Paired Associates test (WMS- IV)
Description	A word-list learning test with built-in cueing. It is a widely used assessment tool to look at explicit, associative episodic memory function. The Verbal Paired Associates (VPA) test is a subtest of the Weschler Memory Scale (WMS), included in its most recent 4th revision (WMS-IV).[1] Word pairs are first read to the patient, followed by one or more recall trials in which the first of the pair is presented to the patient. The patient is asked to provide the other word in the pair. The test exists in many different forms and multiple versions have appeared in the WMS over time.
Functions Tested	Immediate and delayed verbal associative memory, delayed recognition memory
Subscales	No subscale; it is a subtest of the WMS-IV
Number of Items/Scoring	Most commonly 19 word pairs are used, ranging in difficulty from easy (e.g., baby – cries) to hard (e.g., Cabbage – pen). Results include a memory acquisition score, a learning score, delayed recall and recognition scores. Scaled scores range from 1-19. Easier pairs can be separated from harder pairs for separate use in clinical practice, although separate norms not available.
Copyright Status	Copyrighted, Pearson PLC
Administration Time	10 minutes
Normative Psychometric Data	Ample normative exist for healthy adults, stratified in one study for ages 16-54, and ages 55-89 years.[2] Scores are computed as part of the WMS-IV, where they are combined with scores from the Logical Memory Test. There is high test-retest concordance averaging about a one point gain from retesting at 2-12 weeks.[1] VPA in WMS-III had high test-retest stability over many years and showed a practice effect.[3]
Sensitivity and Specificity	Varies across clinical populations. One study, using logistic regression, found that it had an 88-93% correct classification rate for healthy adults and patients with mild AD.[4] Large effect sizes are reported in WMS-IV manual with TBI and AD populations compared to normal adults.
Advantages	It is easy to administer. It has been well studied in multiple, diverse clinical populations, including normal cognitive aging, traumatic brain injury, MCI, AD, Parkinson disease, multiple sclerosis, and patients with alcoholism, among others. There are data which show brain regions activated using fMRI during test completion.[5]

Limitations	It is copyrighted. Normative data for older populations is not always calculated separately for easier and harder paired words. Used by itself, it may not be sufficiently sensitivity to detect Mild Cognitive Impairment.[6, 7] Performance declines with age and is impacted by level of education.
References	1. Wechsler D. Wechsler Memory Scale 4th ed. San Antonio, TX: Pearson, 2008. 2. Uttl B, Graf P, Richter LK. Verbal paired associates tests limits on validity and reliability. Arch Clin Neuropsychol 2002;17:567-581. 3. Lo AH, Humphreys M, Byrne GJ, Pachana NA. Test-retest reliability and practice effects of the Wechsler Memory Scale-III. J Neuropsychol 2012;6:212-231. 4. Lowndes GJ, Saling MM, Ames D, Chiu E, Gonzalez LM, Savage GR. Recall and recognition of verbal paired associates in early Alzheimer's disease. J Int Neuropsychol Soc 2008;14:591-600. 5. Vannest J, Eaton KP, Henkel D, et al. Cortical correlates of self-generation in verbal paired associate learning. Brain Res 2012;1437:104-114. 6. Pike KE, Kinsella GJ, Ong B, et al. Is the WMS-IV verbal paired associates as effective as other memory tasks in discriminating amnestic mild cognitive impairment from normal aging? Clin Neuropsychol 2013;27:908-923. 7. Brooks BL, Iverson GL, Holdnack JA, Feldman HH. Potential for misclassification of mild cognitive impairment: A study of memory scores on the Wechsler Memory Scale-III in healthy older adults. J Int Neuropsychol Soc 2008;14:463-478.

Spatial Cognition

Test Name	Behavioral Inattention Test
Description	A bedside or office-based battery of paper and pencil tasks that assess for spatial cognitive and attentional dysfunction. It was originally designed in 1987[1] to assess unilateral visual neglect, with an aim toward improving rehabilitation strategies. The BIT is made up of two different subtests (see below, subscales), which broadly assess conventional neuropsychological spatial cognitive functions. In the Conventional subtest, patients are asked to perform tasks such as bisecting lines, cancelling letters and star targets in an array, drawing from memory, and copying drawings. In the Behavioral subtest, patients perform tasks such as reading a simulated menu and narrative, sorting coins, and setting the time on a plastic clock.
Specific Functions Assessed	Spatial neglect, visuo-constructional praxis, visuo-perceptual spatial processing
Subscales	2 subscales: Conventional (neuropsychological tasks, including cancellation, drawing, line bisection) and Behavioral (simulated tasks of daily functioning).
Number of Items/Scoring	The Conventional subscale has 6 items, with 146 possible points. The Behavioral subscale has 9 items, with 81 possible points. A combined total score of 227 is possible. Higher scores reflect better performance.
Copyright Status	Copyrighted, Pearson, PLC
Administration Time	40 minutes; 15 minutes (Conventional), 25 minutes (Behavioral)
Normative Psychometric Data	There is limited normative data. One study reports results for 47 healthy controls with an age range of 34-93.[2]
Sensitivity and Specificity	Studies support the validity of the Conventional subtest in detecting functional disability relevant to stroke recovery.[3] No publications reporting sensitive and specificity in other disease groups could be identified. It has been used to detect spatial neglect in AD dementia[4] as well as stroke.[1]
Advantages	Special training background of examiner is not required: it can be performed by an allied health professional, nurse, psychologist or technician. The test has been correlated with functional performance in previous studies. The Conventional subtest has been used alone in many studies and can be administered quickly.
Limitations	Test-retest reliability may be unsatisfactory in acute patients.[5] Scoring

	reflects omission and not commission errors, which the examiner is simply instructed to note; thus, it may not detect ipsilateral neglect or a visual grasp phenomenon that can cause functional disability. It does not distinguish perceptual-attentional ("Where") from motor-intentional ("Aiming") errors. It does not assess personal neglect or far-space deficits. It may not be reliable patients with ocular-based or other non-cognitive visual deficits or aphasia.
References	1. Wilson B, Cockburn J, Halligan P. Development of a behavioral test of visuospatial neglect. Arch Phys Med Rehabil 1987;68:98-102. 2. Stone SP, Halligan PW, Wilson B, Greenwood RJ, Marshall JC. Performance of age-matched controls on a battery of visuo-spatial neglect tests. J Neurol Neurosurg Psychiatry 1991;54:341-344. 3. Katz N, Hartman-Maeir A, Ring H, Soroker N. Functional disability and rehabilitation outcome in right hemisphere damaged patients with and without unilateral spatial neglect. Arch Phys Med Rehabil 1999;80:379-384. 4. Liu CJ, McDowd J, Lin KC. Visuospatial inattention and daily life performance in people with Alzheimer's disease. Am J Occup Ther 2004;58:202-210. 5. Kutlay S, Kucukdeveci AA, Elhan AH, Tennant A. Validation of the Behavioural Inattention Test (BIT) in patients with acquired brain injury in Turkey. Neuropsychol Rehabil 2009;19:461-475.

Test Name	Cancellation Task
Description	This test is a brief test of spatial attention and motor action. The test was popularized by Albert in 1973,[1] although he noted previous, similar tasks. It became a widely used tool to measure aspects of visuo-spatial attention and neglect in patients with acquired brain injury. In most versions, the patient is shown a piece of paper with a clutter of items. Patients are then asked to mark each occurrence of a certain target item, a task which is often timed. The examiner can also gauge patients' planning and organization abilities by noting their strategy for completing the task. Individuals with spatial neglect often miss items on the contra-lesional side of the page.
Specific Functions Assessed	Visuospatial attention for distributed spatial search, fine motor action and coordination
Subscales	No subscales
Number of Items/Scoring	There is a variable number of targets. Several variants exist (line, letter, object, etc.)[2] and tests can be timed or untimed. Most healthy individuals make very few mistakes, with normal cut-offs of 0 to 1 omission. Scoring often includes number of omissions, number of errors (if foils are included), and the location of omissions (left vs. right).
Copyright Status	Public Domain
Administration Time	3 minutes
Normative Psychometric Data	Some normative data exist with several versions for healthy children and adults across the lifespan.[3-5]
Sensitivity and Specificity	There is limited data for any cancellation test version. Healthy individuals generally make 0 to 1 omission on these tests.
Advantages	The test is simple to administer. It may uncover hemi-spatial neglect even in the absence of more overt signs.[6] Its strength as a screening test may be related to its ability to detect a spatial cognitive deficit, even with perseverative behavior or disengagement.
Limitations	There is limited normative data and no standardization of tests, except as part of the Behavioral Inattention Test (BIT).[7] It is not specific to one impaired behavioral network, but may be more sensitive in detecting abnormalities in non-parietal (i.e. frontal or sub-cortical) stroke patients.

References	1. Albert ML. A simple test of visual neglect. Neurology 1973;23:658-664. 2. Gauthier L, Dehaut F, Joanette Y. The Bells Test: A quantitative and qualitative test for visual neglect. J Clin Neuropsychol 1989;11:49-54. 3. Uttl B, Pilkenton-Taylor C. Letter cancellation performance across the adult life span. Clin Neuropsychol 2001;15:521-530. 4. Mark VW, Woods AJ, Ball KK, Roth DL, Mennemeier M. Disorganized search on cancellation is not a consequence of neglect. Neurology 2004;63:78-84. 5. Pradhan B, Nagendra HR. Normative data for the letter-cancellation task in school children. Int J Yoga 2008;1:72-75. 6. Binder J, Marshall R, Lazar R, Benjamin J, Mohr JP. Distinct syndromes of hemineglect. Arch Neurol 1992;49:1187-1194. 7. Halligan PW, Marshall JC. Is neglect (only) lateral? A quadrant analysis of line cancellation. J Clin Exp Neuropsychol 1989;11:793-798.

Test Name	Catherine Bergego Scale
Description	A brief bedside or office-based scale that assesses the presence and severity of spatial cognitive dysfunction and its impact on functional status. It is able to assess deficits in both perceptual- and motor-based behaviors. The CBS was first introduced in 1995[1] and then further characterized[2] as a tool to assess spatial dysfunction in a range of daily activities. Patients are assessed on ten different items in the realms of functional movements, perception and performance. Three administration methods have been used: 1. A therapist-rater observes and immediately scores performance relevant to each of the ten items from 0 (no neglect) to 3 (severe neglect). Although the original publications did not specify how each item should be administered, a controlled, replicable assessment and scoring procedure for this method is available.[3] 2. The items can be scored off-line by a therapist (0-3 for each item as above) who is familiar with the patient's clinical status. This is based on this therapist's estimate of a patient's average performance relevant to each item over a period of time. 3. The scale can be administered as a questionnaire read by or to the patient, who rates self-performance from 0-3 on each of the ten items, as above. This method is used to assess self-awareness of spatial neglect rather than to assess spatial performance directly.
Specific Functions Assessed	Overt performance of functional-related tasks, functional consequences of spatial cognitive deficits, integrative functioning of perceptual, representational, and motor domains
Subscales	Includes a perceptual-attentional (PA) and motor-exploratory (ME) subscale, based on a principal component analysis.[4] Both subscales are normally administered.
Number of Items/Scoring	10 items; the maximum (worst) score is 30
Copyright Status	Public Domain; Kessler Foundation has proprietary training process for reliable administration (KF-NAP)™
Administration Time	15 minutes
Normative Psychometric Data	Normative data is available only for populations of stroke patients.[5] Healthy individuals should obtain a score of 0.
Sensitivity and Specificity	It ranges from 76-94% sensitivity in detecting spatial neglect in stroke

	patients.[3-7] It has been studied only in stroke patients to this point.
Advantages	It is compatible with clinical care evaluations. It can be performed by allied health professionals, like nurses, psychologists or technicians. Special reliability training has been recommended.[3] The test has value in identifying different spatial neglect syndromes. It is one of few spatial tests which could, in a self-rating format, assess anosoagnosia. It has a particular strengths in assessing performance in personal (body parts or on the body surface), peri-personal (within arm's reach) and extra-personal spaces (beyond arm's reach).
Limitations	The test is more sensitive to omission than commission errors (i.e., it may be less sensitive in detecting visual grasp and ipsilateral neglect). The range of reliability in clinicians who have not completed training may be unsatisfactory. It may not be able to differentiate whether sensory neglect or motor neglect contribute to the observed functional difficulties.
References	1. Bergego C, Azouvi P, Samuel C, et al. Validation d'une échelle d'évaluation fonctionelle de l'héminégligence dans la vie quotidienne: L'échelle c.B. Annales de Réadaptation et de médecine physique 1995;38:183-189. 2. Azouvi P. Functional consequences and awareness of unilateral neglect: Study of an evaluation scale. Neuropsychol Rehabil 1996;6:133-150. 3. Chen P, Hreha K, Fortis P, Goedert KM, Barrett AM. Functional assessment of spatial neglect: A review of the Catherine Bergego Scale and an introduction of the Kessler foundation neglect assessment process. Top Stroke Rehabil 2012;19:423-435. 4. Goedert KM, Chen P, Botticello A, Masmela JR, Adler U, Barrett AM. Psychometric evaluation of neglect assessment reveals motor-exploratory predictor of functional disability in acute-stage spatial neglect. Arch Phys Med Rehabil 2012;93:137-142. 5. Azouvi P, Olivier S, de Montety G, Samuel C, Louis-Dreyfus A, Tesio L. Behavioral assessment of unilateral neglect: Study of the psychometric properties of the Catherine Bergego Scale. Arch Phys Med Rehabil 2003;84:51-57. 6. Azouvi P, Samuel C, Louis-Dreyfus A, et al. Sensitivity of clinical and behavioural tests of spatial neglect after right hemisphere stroke. J Neurol Neurosurg Psychiatry 2002;73:160-166. 7. Luukkainen-Markkula R, Tarkka IM, Pitkanen K, Sivenius J, Hamalainen H. Comparison of the Behavioural Inattention Test and the Catherine Bergego Scale in assessment of hemispatial neglect. Neuropsychol Rehabil 2011;21:103-116.

Test Name	Clock Drawing Test
Description	The Clock Drawing test is a simple, "bedside" test that assesses various executive and spatial cognitive functions. The test originated in the 1950s, promoted by Critchley[1] and others, to detect parietal lobe dysfunction. More recently, it has been used as a screening tool for constructional apraxia, common in certain dementias and other neurologic disorders.[2] In most versions, the patient is shown a blank circle and asked to fill in the numbers of a clock-face and 2 hands demonstrating a specific time (e.g., 10 min. past 11). Sometimes, the paper is blank, without the pre-drawn circle. A popular version called CLOX[3] does include a second part as well. In that version, Part 1 is given first. Then, in part 2, the patient observes the examiner drawing a complete detailed clock-face with the same specified time. The patient is then asked to copy the figure in an adjacent space while still viewing the figure. This version was designed to try detect and discriminate more "pure executive" dysfunction from any non-executive constructional apraxia. The score difference between parts 1 and 2 is thought to reflect the specific contribution of executive control versus construction praxis in completing the task (see Scoring below).
Specific Functions Assessed	Executive functions, including planning, organization, simultaneous processing, and visuo-constructional praxis (directly); Aspects of attention, semantic memory, auditory language, and grapho-motor functions (indirectly)
Subscales	The test has 1 or 2 parts (see Description). The most commonly used component involves instruction to draw a clock and all of its components, with hands set at a specific time. A second component can be used to help discriminate executive function from constructional praxis; the patient observes the examiner drawing the clock and is then asked to copy the clock separately.
Number of Items/Scoring	Many scoring systems are used and have been studied. These incorporate qualitative and quantitative aspects of three main determinants: integrity of the clock's face, sequence and presence of the numbers within the clock, and the presence and placement of the clock's hands.[4] A 6-point error score is widely used, where 6= "no reasonable representation of a clock" and 1= "perfect". A 20-Item Clock Drawing Interpretation Scale has been studied as a dementia screening scoring system; it has an "abnormal" cut-off of < 19 points. A more recent 7-point scale has 3 sub-scales: time (3 points, two hands depict accurate time), numbers (2 points, numbers are all

	present and inside clock), and spacing (2 points, numbers evenly spaced from each other/circle edge).[5]
Copyright Status	Public Domain
Administration Time	Approximately 3-7 minutes (varies according to patient and examiner)
Normative Psychometric Data	Normative data is available primarily for adults age 40s to 90s; most of these studies use the 20-item and 7-point scoring systems (see Scoring above).[6, 7]
Sensitivity and Specificity	Sensitivity and specificity are variable based on dementia type, but each have ranges from 75-92%. Patient populations that have been studied include various degenerative dementias, delirium, psychological disorders, and cerebrovascular disease.[8]
Advantages	The test is short, simple, well-tolerated, easy to administer, inexpensive, and able to evaluate various cognitive functions. Culture, gender, and education do not seem to affect the test. It is a good initial and longitudinal dementia screen, especially for older adults. It is a reasonably reliable tool for executive function screening as it correlates with other domain-specific cognitive tests.[9]
Limitations	The variety of methods of scoring and administration make normative values difficult to interpret. Performance on the test may be affected by the existence of focal motor deficits or visual neglect.
References	1. Critchley M. The parietal lobes. New York: Hafner, 1953. 2. Lezak MD, Howieson DB, D.W. L. Neuropsychological assessment, 4th ed. ed. New York: Oxford University Press, 2004. 3. Royall DR. Comments on the executive control of clock-drawing. J Am Geriatr Soc 1996;44:218-219. 4. Mendez MF, Ala T, Underwood KL. Development of scoring criteria for the clock drawing task in Alzheimer's disease. J Am Geriatr Soc 1992;40:1095-1099. 5. Freund B, Gravenstein S, Ferris R, Burke BL, Shaheen E. Drawing clocks and driving cars. J Gen Intern Med 2005;20:240-244. 6. Nyborn JA, Himali JJ, Beiser AS, et al. The Framingham Heart Study clock drawing performance: Normative data from the offspring cohort. Exp Aging Res 2013;39:80-108. 7. Hubbard EJ, Santini V, Blankevoort CG, et al. Clock drawing performance in cognitively normal elderly. Arch Clin Neuropsychol 2008;23:295-327. 8. Adunsky A, Fleissig Y, Levenkrohn S, Arad M, Noy S. Clock drawing task, mini-mental state examination and cognitive-functional independence measure: Relation to functional outcome of stroke patients. Arch Gerontol Geriatr

	2002;35:153-160. 9. Lorentz WJ, Scanlan JM, Borson S. Brief screening tests for dementia. Can J Psychiatry 2002;47:723-733.

Test Name	Cube Copying Test
Description	This test is a widely used bedside test of visuospatial function, planning, and constructional skills. It has been included as the "visuo-constructional" component of several widely used global cognitive screening tools, including the neuropsychology assessment of the Consortium to Establish a Registry for Alzheimer Disease (CERAD), the Short Test of Mental Status (STMS), and the Montreal Cognitive Assessment (MoCA). In this test, the patient views a 3-dimensional cube and is asked to copy it in an adjacent space.
Specific Functions Assessed	Visuospatial, executive, and constructional praxis
Subscales	None
Number of Items/Scoring	It is typically scored as correct or incorrect. It is often part of a global battery (like the STMS or MoCA). Variable quantitative scoring methods do exist using the number of correct connections, planes, and features, such as 3-demensional perspective.
Copyright Status	Public Domain
Administration Time	2 minutes
Normative Psychometric Data	Ample normative data are available for healthy elderly adults (mean ages 67-72) and patient with degenerative dementia.[1-3]
Sensitivity and Specificity	Reliability as a measure that is separate from a larger test battery has not been well validated. The test has been studied mostly in clinical populations with stroke[4] and AD dementia.[2]
Advantages	The test is simple to administer. It is sensitive to impairment in multiple neurodegenerative diseases and hemi-spatial neglect.
Limitations	The test is nonspecific and minor errors may be frequent in normal aging. It may be limited in its screening potential alone rather than as part of a test battery.
References	1. Plude DJ, Milberg WP, Cerella J. Age differences in depicting and perceiving tridimensionality in simple line drawings. Exp Aging Res 1986;12:221-225. 2. Welsh KA, Butters N, Mohs RC, et al. The Consortium to Establish a Registry for Alzheimer's Disease (CERAD). Part V. A normative study of the neuropsychological battery. Neurology 1994;44:609-614. 3. Maeshima S, Osawa A, Maeshima E, et al. Usefulness of a cube-copying test in outpatients with dementia. Brain Inj 2004;18:889-898. 4. Arrigoni G, De Renzi E. Constructional apraxia and hemispheric

	locus of lesion. Cortex 1964;1:170-197.

Test Name	Benton Judgment of Line Orientation (JLO)
Description	A spatial cognitive test designed to assess aspects of basic visuospatial perception and judgment. It was first described by Benton et al in 1975[1] for the purpose of detecting right hemisphere dysfunction in patients with stroke and other acquired injuries. Patients are asked to judge a pair of angled lines that visually match an identical pair that is immersed within a semi-circular array of 11 lines. Patients are asked to indicate which two lines from the array are in exactly the same position and point in the same direction as the pair of lines in question (separated from the array). Because it assesses a basic level of visuospatial ability, JLO has enjoyed widespread use in neuropsychological assessment, in many different populations. Because it is seen as a tool to assess rudimentary visuospatial ability, it can useful in interpreting a patient's performance on more complex tasks of visual reasoning and visuo-construction.
Specific Functions Assessed	Perception of visuospatial relations
Subscales	None, but various short forms where half the items are available
Number of Items/Scoring	5 practice items and 30 scored items; max (best) score is 30, corrected for age and gender
Copyright Status	Copyright, Psychological Assessment Resources, Inc. (PAR)
Administration Time	20 minutes for the full version; 10 minutes for various short-forms; scoring takes about 5 minutes
Normative Psychometric Data	Ample normative data is available for healthy individuals, mostly age 50s-80s.[2] Normative data for short forms are also available.[3]
Sensitivity and Specificity	The sensitivity and specificity is good. The test also has good internal consistency and test-re-test reliability. It has been studied mostly in patients with stroke (comparing right vs. left hemisphere lesions)[1], Parkinson disease (PD), and AD dementia,[2] including the posterior cortical atrophy (PCA) variant. Short forms may be reliable in detecting spatial deficits in Parkinson disease (PD) patients.[4]
Advantages	It is simple to administer. It may be of the "purest" spatial cognition tests available, perhaps requiring only simple attention and spatial networks.
Limitations	The test may not discriminate between different dementia etiologies. It may be selective for perceptual-representational dysfunction and not identify deficits in motor-intentional or visual-exploratory function. It may not be sensitive to right hemisphere pathologies. If there is inherent hemi-spatial bias, it may not detect an abnormality

	in angular perception.
References	1. Benton A, Hannay HJ, Varney NR. Visual perception of line direction in patients with unilateral brain disease. Neurology 1975;25:907-910. 2. Ska B, Poissant A, Joanette Y. Line orientation judgment in normal elderly and subjects with dementia of Alzheimer's type. J Clin Exp Neuropsychol 1990;12:695-702. 3. Qualls CE, Bliwise NG, Stringer AY. Short forms of the Benton Judgment of Line Orientation Test: Development and psychometric properties. Arch Clin Neuropsychol 2000;15:159-163. 4. Gullett JM, Price CC, Nguyen P, Okun MS, Bauer RM, Bowers D. Reliability of three benton Judgment of Line Orientation short forms in idiopathic Parkinson's disease. Clin Neuropsychol 2013;27:1167-1178.

Test Name	Line bisection test
Description	This test is a brief bedside/office test that assesses the presence and severity of spatial bias. It was popularized as a multi-step trial test in 1980 by Schenkenberg et al[1] to detect lateralized hemispheric dysfunction in patients with brain injury. It has subsequently had widespread use in many patient populations to detect spatial bias and hemi-spatial neglect. In the test, the patient is asked to mark the center of a long (> 20 cm) horizontal line on a blank page.
Specific Functions Assessed	Two-dimensional spatial computation, horizontal spatial bias
Subscales	No subscales
Number of Items/Scoring	Different methods are recommended. Generally, 5-10 trials are recommended for a reliable result. Scoring typically involves marking the direction and mean deviation from midpoint of the line.
Copyright Status	Public Domain
Administration Time	1-2 minutes
Normative Psychometric Data	Some normative data are available for healthy adults, including the elderly, with mean age in 70s.[2]
Sensitivity and Specificity	The sensitivity of the test alone to detect spatial bias may be near 76%. Its sensitivity, however, is frequently criticized, as many patients with deficits are relatively accurate. Test-retest reliability is good in several studies.[3] Some reports indicate that a rightward bias >15mm may be specific for right hemisphere dysfunction, particular in women (of all ages); this rightward bias seems less specific for older men.[4] It has been studied mostly in patients with stroke.[5]
Advantages	It provides a rapid, quantitative assessment, which can be appropriate for even severely impaired patients. It can be modified with cues and distracters in order to separately identify left-sided unawareness and leftward directional hypokinesia.[6] A video-adapted version can also separate “where” and “aiming” biases. Accurate performance predicted better functional stroke outcomes in one study.[7]
Limitations	It is not maximally sensitive when used alone as opposed to part of a battery. Video-adapted and other automated versions are not commercially available.

References	1. Schenkenberg T, Bradford DC, Ajax ET. Line bisection and unilateral visual neglect in patients with neurologic impairment. Neurology 1980;30:509-517. 2. Barrett AM, Craver-Lemley CE. Is it what you see, or how you say it? Spatial bias in young and aged subjects. J Int Neuropsychol Soc 2008;14:562-570. 3. Bailey MJ, Riddoch MJ, Chrome P. Test-retest stability of three tests for unilateral visual neglect in patients with stroke: Star cancellation, line bisection, and the baking tray task. Neuropsychol Rehabil 2004;14:403-419. 4. Chen P, Goedert KM, Murray E, Kelly K, Ahmeti S, Barrett AM. Spatial bias and right hemisphere function: Sex-specific changes with aging. J Int Neuropsychol Soc 2011:1-8. 5. Sea M-JC, Henderson A. The reliability and validity of visuospatial inattention tests with stroke patients. Occup Ther Int 1994;1:36-48. 6. Schwartz RL, Barrett AM, Kim M, Heilman KM. Ipsilesional intentional neglect and the effect of cueing. Neurology 1999;53:2017-2022. 7. Friedman PJ. Spatial neglect in acute stroke: The line bisection test. Scand J Rehabil Med 1990;22:101-106.

Test Name	Navon Figures Test
Description	This test is a brief test of hierarchical visual processing that resolves global versus local feature precedence. It was originally popularized by David Navon[1] in 1977 to test a hypothesis called *global advantage* in perceptual processing. Global advantage it is the notion that we normally process global perceptual features before local ones. It has been used widely to look for evidence of simultanagnosia and other higher-order, visuo-perceptual deficits in various neurological disorders. The test consists of showing patients simple figures of several large shapes ("Compound figures"), which are composed of many smaller, usually different, shapes. The patient is asked to describe what is shown in the
Specific Functions Assessed	Visual hierarchical processing bias (i.e., detail and holistic visual perception biases), mental flexibility related to set-shifting
Subscales	None
Number of Items/Scoring	A variable number of figures can be used. It is scored as to whether the proper target, based on instruction, is identified (i.e. the "small" or "large" shape). There are no accepted standardized scoring schema.
Copyright Status	Public Domain
Administration Time	3-5 Minutes
Normative Psychometric Data	Sparse normative data exist for healthy individuals, across the lifespan. Some of this data is based on reaction time rather than identification errors (as little or no errors are made by healthy adults).[2, 3]
Sensitivity and Specificity	It has been studied in patients with autism,[4] obsessive compulsive disorder (OCD),[5] Parkinson disease,[6] and developmental dyslexia.[7]
Advantages	It is simple to administer. It may unmask differential processing in "global" versus "local" visual attentional processes, problems with set-shifting or atypical dominance of brain hemispheres. It is one of a set of simple tasks that can be used to detect simultanagnosia.
Limitations	There is no standardized scoring system or normative data to judge performance. If interested in reaction time instead of errors, it is impractical for bedside testing because of millisecond (ms) temporal resolution. Global precedence in performing this task may be less common in young children and those older than 50.[3] It may not be valid for assessment in non-Western cultures.[8]

References	1. Navon D. Forest before trees: The precedence of global features in visual perception. Cogn Psychol 1977;9:353-383. 2. Lamb MR, Robertson LC. Do response time advantage and interference reflect the order of processing of global- and local-level information? Percept Psychophys 1989;46:254-258. 3. Staudinger MR, Fink GR, Mackay CE, Lux S. Gestalt perception and the decline of global precedence in older subjects. Cortex 2011;47:854-862. 4. Wang L, Mottron L, Peng D, Berthiaume C, Dawson M. Local bias and local-to-global interference without global deficit: A robust finding in autism under various conditions of attention, exposure time, and visual angle. Cogn Neuropsychol 2007;24:550-574. 5. Moritz S, Wendt M, Jelinek L, Ruhe C, Arzola GM. No disadvantage for the processing of global visual features in obsessive-compulsive disorder. J Int Neuropsychol Soc 2008;14:489-493. 6. Barrett AM, Crucian GP, Schwartz R, Nallamshetty H, Heilman KM. Seeing trees but not the forest: Limited perception of large configurations in PD. Neurology 2001;56:724-729. 7. Rubinsten O, Henik A. Double dissociation of functions in developmental dyslexia and dyscalculia. J Educ Psychol 2006;98:854-867. 8. Davidoff J, Fonteneau E, Fagot J. Local and global processing: Observations from a remote culture. Cognition 2008;108:702-709.

Test Name	Rey-Osterrieth Complex Figures Test (ROCFT)
Description	This test is a relatively brief spatial cognitive and executive function test that requires the immediate and delayed copy of a complex figure. It was first developed in the 1940s as a tool to detect deficits in adults with brain trauma. It was then adapted to study "gestalt" knowledge within the normal neurodevelopmental stages of childhood.[1, 2] It has become one of the most widely used neuropsychological measures in neuropsychiatric disorders, including dementias, developmental learning disabilities, and disorders, such as OCD and schizophrenia.[3] In the ROCFT, a patient is shown a complex figure and asked to copy the figure immediately. Without forewarning, the patient is then asked to reproduce the figure, from memory, after a short-delay of typically 3 minutes. The patient is asked again to reproduce the figure again after a long delay of 30 minutes. Some later versions supplemented the ROCFT by adding recognition and matching trials following the 30-minute delayed recall.[4, 5]
Specific Functions Assessed	Visuo-constructional praxis, visual-spatial perception, planning and organization (executive), working memory, episodic visual memory
Subscales	No subscales
Scoring	A number of explicit methods for manual or computerized scoring are available. The Boston Qualitative Scoring System (BQSS) is a comprehensive scoring system that provides 17 qualitative scores and 6 summary scores.[6]
Copyright Status	Public domain[8]
Administration Time	Approximately 25-30 minutes, in addition to a 30-minute delay period
Normative Psychometric Data	Ample normative values are available, mostly for healthy individuals aged 30 to 85.[4] Additional norms are available for other ages and have been adjusted for some demographic factors.[7, 8]
Sensitivity and Specificity	Vary across clinical populations and studies. It has been well-studied in AD and other dementias, schizophrenia, and obsessive compulsive disorder, among other disorders.
Advantages	The test is easy to administer. An explicit scoring system with associated norms is available. The task is sensitive to a variety of cognitive and perceptual deficits and can detect problems within many domains, including memory, spatial cognition, and executive function.

[8] Some scoring systems are copyrighted (e.g., Boston Qualitative Scoring System, by Psychological Assessment Resources [PAR]).

Limitations	Patients can make errors for many reasons (e.g., perceptual difficulties, visuo-construction problems, episodic memory), which limits the specificity of the test.
References	1. Waber DP, Holmes JM. Assessing children's memory productions of the Rey-Osterrieth Complex Figure. J Clin Exp Neuropsychol 1986;8:563-580. 2. Watanabe K, Ogino T, Nakano K, et al. The Rey-Osterrieth Complex Figure as a measure of executive function in childhood. Brain Dev 2005;27:564-569. 3. Spreen O, Strauss E. A compendium of neuropsychological tests. New York: Oxford University press, 1998. 4. Fastenau PS, Denburg NL, Hufford BJ. Adult norms for the Rey-Osterrieth Complex Figure Test and for supplemental recognition and matching trials from the Extended Complex Figure Test. Clin Neuropsychol 1999;13:30-47. 5. Fastenau PS. Development and preliminary standardization of the "extended complex figure test" (ECFT). J Clin Exp Neuropsychol 1996;18:63-76. 6. Stern RA, Javorsky DJ, Singer ES, et al. The Boston Qualitative Scoring System for the Rey-Osterrieth Complex Figure: BQSS : Professional manual: Psychological Assessment Resources, 1999. 7. Duley JF, Wilkins JW, Hamby SL, Hopkins DG, Burwell RD, Barry NS. Explicit scoring criteria for the Rey-Osterrieth and Taylor complex figures. Clinical Neurologist 1993;7:29-38. 8. Gallagher C, Burke T. Age, gender and IQ effects on the Rey-Osterrieth Complex Figure Test. Br J Clin Psychol 2007;46:35-45.