Exemplary Tests	Psychometrics	Clinical Sensitivity	Normative data	Advantages	Disadvantages
CLASSICAL TEST BATTERIES (with focus on gestures)					
The Florida Apraxia Battery – Extended & Revised Sydney (FABERS) [1] Most items selected from existing Battery (Florida Apraxia Battery – FAB) [35] and other public measures. Detailed assessment battery: pantomime recognition, pantomime discrimination, action semantic knowledge, verbal-visual semantic knowledge, and pantomime expression. Time to administer: 45 minutes [43] Scoring (videotaped): 1 point for each correct item/compound; 0 points for item performed with any errors.	Interrater-Reliability (videotaped): ≥ 89% Inter-correlations: Strong correlations between transitive pantomime tasks: ρ = .64 (photograph – command) ρ = .73 (photograph - imitation) ρ = .52 (command - imitation) Tool naming correlates with transitive pantomime to command: ρ =.72 Correlation tool selection task - transitive pantomime-to-photograph task: ρ = .54 Correlation meaningless imitation task - pantomime recognition task: ρ = –.65 **[1]**		Right-handed healthy participants from the Syd-ney Metropolitan Region in Australia (N = 16, 55 – 83) **[1]** Means, standard deviation, ranges, and Kolmogorov–Smirnov *Z* statistics on the tests for 16 healthy participants are provided **[1]**	Distinguishes transitive from intransitive pantomimes, and meaningless from meaningful pantomimes.	Relatively small number of healthy participants. Ceiling effects in some subtests.
Test of Upper Limb Apraxia (TULIA) [2] Six subtests for both imitation and pantomime of non-symbolic, intransitive and transitive gestures. Test comprised 48 items compiled from various sources. Time to administer: 20 minutes Scoring (videotaped): 6-point scoring method (min. score = 0; max. score = 5 for each item) 1) Achievement of overall movement: score 0 or 1 vs. 2 – 5 2) detailed error analysis (e.g. body part as object errors) for final scoring (2 – 5).	Criterion validity: Significantly lower TULIA scores of patients with left brain damage compared to right brain damage. Construct validity: strong correlation between TULIA and De Renzi test [3] (r = .82). Interrater reliability (videotaped): .65 - .99 (kappa) Retest reliability: α = .83 (examined three times within 24 h) **[2]**	Cut-off score = 194 (two SD below mean score of healthy subjects). Within the LBD group 13% patients were severely apraxic (<65), 12% moderately (<130) and 43% mildly apraxic [4]. In the RBD, 2% were moderately and 37% mildly apraxic. Patients with cortical lesions were more frequently apraxic, particularly if severely (10.4% vs. 2.7%) or moderately affected (9.4% vs. 5.4%) **[2]**	Normative data: 50 healthy subjects originating from Switzerland and Germany (age: 43–93); 84 subjects with left brain damage (age: 30–84), 49 subjects with right brain damage (age: 33–85), **[2]**	Short time to administer; high reliability and validity; detailed scoring system. Materials published in English: **[2]**	Imitation of nonsymbolic gestures in part less reliable. **[2]**

De Renzi Test [3] Requires patients to imitate 24 gestures. Test entails symbolic and nonsymbolic movements, positions and sequences, gestures that involve independent finger movement and gestures that involve movement of the whole hand. Time to administer: 15 minutes Scoring (non-videotaped): 3, 2, 1 or 0 points (based on number of trials needed)	Correlation of total score with demonstration-of-use test (participant asked to pantomime use of ten common tools) = .80	Max. score: 72 Distribution of scores in controls: 72 – 62 points →cutoff score: < 62 points 20% of RBD and 50% of LBD patients scored under cutoff	Normative data: 100 control patients without brain injury tested in Italy (mean age: 52.6); 80 RBD (mean age: 57.7); 100 LBD (mean age: 55.9). Patients carried out task with ipsilateral side of lesion. Half of controls performed test with left limb, other half with right limb.	Easy to administer. Simple scoring system. Feasible for aphasic patients, too. Materials described in English: [3]	No psychometric properties available.
De Renzi "ideational" Test [48] Demonstration of tool use: mixture of pantomime and real use: grasp the object and show how to use (e.g. with hammer in hand but without recipient object, i.e. no nail). Entails an "ideational apraxia" subtest and an "ideo-motor" apraxia subtest. Time to administer: -	External validity: Rho = .410 (Corr. with ideo-motor subtest) Rho = .392 (Corr. with Raven´s progressive matrices [49]) Rho = -.265 (Corr. with visual reaction times [50]) Rho = .579 (Corr. with Aphasia comprehension score)	IA subtest: Cutoff-score: > 14 points (total score: 14 points). Errorless performance of all RBD patients and all control patients. 45 LBD patients made ≥ 1 error. IMA subtest: Cutoff score: > 17 points (total score: 20). No control/ RBD scored under cutoff. 45/160 LBD scored under cutoff.	40 control patients without cerebral lesions tested in Italy, 45 RBD patients and 160 LBD patients	Tasks with tools appear easier to instruct: advantage when testing severely aphasic patients. Materials described in English: [48]	Test includes gestures that are culture-specific to Italy. Ceiling effects especially for IA subtest.
Test of Bartolo et al. [5] Adapted version of De Renzi Test. Tasks including gesture production (intransitive gestures, gestures including real objects, meaningless gestures) and gesture recognition tasks (pantomimes and intransitive gestures). Tasks were based on verbal command, visual modality, imitation modality, object use and tactile modality (touching object before pantomiming its use). Time to administer: 2 h (in healthy participants) [43]	Items eliciting least errors in healthy subjects were selected. **[5]**	Means of healthy subjects for each subtest were close to ceiling or at ceiling. **[5]** Range, means, standard deviations and cut-off scores are reported for each subtest. Production of intransitive gestures Verbal (15.95/16) and visual (15.5/16). Production of Pantomimes: Verbal (15.6/16), visual (15.7/16) and tactile (15.6/16). Recognition: intransitive gestures (39.2/16) and pantomimes (39.0/40). Remaining tasks performed at ceiling: transitive gestures	Normative data: 60 British subjects (age: 17–81); healthy, right-handed subjects.	21 tasks including real objects, videotaped stimuli available for some subtests. Materials described in English: **[5]**	Stimuli of the identification and the production tasks are the same. For reducing learning effects, test needs to be administered on two consecutive days. Test battery hasn´t been applied to patient sample so far. Ceiling effects. No interrater reliability reported.

Scoring (non-videotaped): 1 point for each correctly executed task (allows for repeated or corrected responses).		(16=16), imitation of intransitive gestures (16=16), pantomimes (16=16), meaningless gestures (20=20), identification of intransitive gestures (16=16) and pantomimes (16=16). **[5]**			Control group shows a broad age range. Within the tested group, no significant correlation between age and task performance was detected. ($p > 0.1$).
BATTERIES AND TESTS INCLUDING NATURALISTIC SINGLE- AND MULTI-STEP TASKS WITH TOOLS AND OBJECTS					
Diagnostic Instrument for Limb Apraxia - Short Version (DILA-S) [7] Shortened version of the DILA – most sensitive items were selected. 3 classic subtests: imitation of meaningless and meaningful gestures, pantomime of tool use, 3 subtests with tools and objects: selection and application of novel tools (NTT) and familiar tools (FTT) and a familiar multistep naturalistic action task (NAT): Preparing Breakfast Time to administer: in total about 40 minutes (typical range: 25-65 minutes), subtests range from 2-15 minutes Scoring (non-videotaped; however, experience needed): different levels of scoring, including criteria driven production scores	Interrater reliability (without videotape by 1st examiner; with videotape by 2nd examiner): (percentage of agreement): Imitation of meaningful gestures: 93%; imitation of meaningless gestures: 97%; pantomime: 93%; Novel Tools: 95%; Familiar Tools: 98%. Internal consistency: meaningless gestures *CR (Composite Reliability) = .772*; meaningful gestures CR = .549; pantomime CR ≥ .884; Familiar Tools CR ≥ .768; Novel Tools CR ≥ .742 Content validity: *Pantomime τ ≥ .634; Familiar Tools Test τ ≥ .471; Novel Tools Test τ ≥ .535* External validity: Correlation with AST: τ = .500; correlation with KAS [10]: τ = .522 **[7].**	For subtests demonstrating differences between subgroups age-specific-cutoff-scores (21 – 50 vs. 51 – 80) and gender-specific cutoff-scores are reported in the manual. Based on patient data, differentiations are possible between mild, moderate and severe apraxia **[7]**. Apraxic symptoms in at least one subscale were shown in 37/44 LBD patients, 7/26 MS-patients, 20/44 TBI patients, 24/27 dementia patients. **[8]**	Normative sample: German sample, 82 right-handed healthy subjects tested uni-manually (age: 21 – 80); Psychometrics based on: 33 LBD subjects (age: 30 – 79); 20 RBD subjects (age: 27 – 78) **[7]**	Classic subtests + includes real tools and naturalistic actions. Single subtests can be evaluated. Free materials, manual and video examples available in: English and German: https://www.moco.uni-konstanz.de/publikationen/assessments/	Subtest Novel Tools Test: novel tools and objects need to be manufactured (note: CAD-files available upon email request). Testing and scoring at the same time requires high familiarity with the material and experience with evaluating apraxic behavior. The differentiation of normative age and gender bins in some subtests leads to reduced numbers of participants per bin.
Naturalistic action Test (NAT) [6] Measures everyday action impairment via 3 settings requiring multiple steps with real objects (e.g. wrapping a gift). Scoring (non-videotaped): Each item scored for accomplishment and errors	Interrater reliability (without videotape by 1st examiner; with videotape by 2nd examiner): 0.98 (kappa) Internal consistency: α = 0.753 (for patient sample); α = 0.793 (for entire sample) Concurrent validity: significant corelations between NAT and FIM	NAT scores: Patients = 10.9/18 (SD 5.5; TBI = 12.6 RBD = 10.3, and LBD 10.3) and controls = 17.3/18 (SD 1.2) (U = 355, $p < .001$) **[6]**	Normative sample: 28 non-neurological control subjects tested in the USA (age: 18 – 80) Psychometrics based on: 30 LBD subjects (age: 34–80), 45 RBD subjects (age: 37–80), 25 subjects	psychometric properties; test includes several multistep naturalistic tasks with real objects. A lateralized attention score is included to allow the	Requires multiple objects and storing. Testing and scoring at the same time requires high familiarity with the material and experience with

for necessary steps. Accomplishment and error scores are combined into a 6 point scale.	cognitive [36], ranged from 0.511 (RBD) to 0.718 (TBI) Construct validity: significant correlations of all items with Star Cancellation Test [37] and DT Baseline sub-test. Predictive validity: NAT and IADL scores for the subjects in the 6-month follow-up were statistically significant (r = .58, N = 48) and better in prediction than FIM cognitive **[6].**		with traumatic brain injury (TBI) (age: 18–66), **[6]**	evaluation of visuo-spatial impairments, free English manual available at: https://mrri.org/naturalistic-action-test/	evaluating apraxic behavior.
LIMB APRAXIA SCREENINGS:					
Apraxia Screen of TULIA (AST) [9, 44] 12 items of TULIA were selected based on an item reduction. Time to administer: 3 minutes Scoring (non-videotaped): 1 = pass; 2 = fail (if major spatial/ semantic errors occur)	Correlation between AST and TULIA: r = 0.96 [9] Test-retest reliability (within 24h): r = 0.95 [9] Specificity = 100%; Sensitivity = 95% (validated with TULIA) [9] Discriminative validity: no association with MDS-UPDRS part III, [45] (measuring parkinsonian symptoms; r = 0.17, p = 0.1) [44] Convergent validity: association with MDS-UPDRS part II [45] (measuring ADL; r = -0.35) and Hoehn & Yahr stage (r = -0.32). [44]	Cutoff-scores: 9 = mild apraxia; 5 = severe apraxia (12 = max. score). Of the entire sample (N = 32), 12 were diagnosed as mildly apraxic and 7 were diagnosed as severely apraxic. **[9]**	Validation sample: 5 RBD and 26 LBD stroke patients (average age: 63.4). All subjects were right-handed. **[9]**	Short time to administer. Materials described in English [9] and German: [51]	Small sample size used for validation.
Kölner Apraxie Screening (KAS) [10, 43] Patients have to imitate abstract and symbolic gestures (buccofacial- and hands-/arms-items) as well as to pantomime use of objects. For all tasks, photos of objects/ gestures are presented as stimulus material. Scoring (non-videotaped): 4 – 0 points for each item (depending in item-specific scoring criteria) Time to administer: 10 Minutes	Internal consistency: α = .968 Homogeneity: .604 Interrater reliability (videotaped): rho = .907 Construct validity: .653 and .565 with two subtests of an imitation test of Goldenberg [47]	Cut-off: ≤ 76 (total score = 80) Sensitivity = 80% Specificity = 98% Mean test scores of healthy subjects = 79.5 (SD = 1.2); LBD patients without apraxia = 78.1 (SD = 3.8); LBD patients with apraxia = 57.6 (SD = 24.8).	Normative sample: German sample, 48 healthy subjects (age: 32 – 73) 106 LBD patients without apraxia (age: 21 - 89) 80 LBD patients with apraxia (age: 25 - 90)	German materials purchasable: https://www.testzentrale.de/shop/koelner-apraxie-screening.html	

Short apraxia screening test (SAST) [11] Consists of 12 items including: intransitive gestures, transitive gestures elicited under verbal, visual, and tactile modalities, imitation of meaningful and meaningless postures and movements, and a multiple object test. Scoring (videotaped): counting of errors	AUC (area under the curve) for apraxia diagnosis, when compared to FAST–R [39], was .928. Interrater reliability (videotaped): .918 (Cohen's kappa) **[11]**	Cutoff-score: >2 errors. Sensitivity = 92%; Specificity = 79%. All patients had IMA (exhibited spatiotemporal errors in particular when performing transitive gestures to verbal commands). 8 patients had CA (committed unrelated content errors). 2 patients with large lesions in the postsylvian region failed the multiple object test. No patient exhibited LKA. **[11]**	Normative sample: 40 healthy controls. 70 LBD patients, All subjects were right-handed originating from Argentina. **[11]**	High sensitivity and specificity. Captures all types of upper limb praxic deficits. Materials described in: English **[11]** Includes online material and videos in the supplemental material: https://www.tandfonline.com/doi/full/10.1080/13803395.2014.951315	
Short test for ideo-motor apraxia (STIMA) [12] Imitation of 18 known and 18 new gestures. Half of the known and new gestures were performed mainly with the proximal part of the upper limb and the remaining half with the distal part. Recognition task of the 18 known gestures afterwards. Time to administer: 4–5 min (when administered to apraxic person) Scoring (videotaped): 0 – 2 points for each item (based on number of trials needed)	Interrater reliability (videotaped): Only one subject of 112 subjects excluded due to different scorings between judges. **[12]**	Recognition of known gestures: 100 %. Age-specific equivalent scores and percentiles for all 8 subtests are reported in the supplementary material. Cutoff-score: <5th percentile. **[12]**	Normative sample: Consists of 111 healthy, right-handed participants (age: 30–84) originating from Italy **[12]**	Usable for bedside screening. Differential evaluation of the body segments (proximal vs. distal). Materials described in: English [12]	No real objects included.
Sensitive Gestures for Screening Apraxia (SFA) [13] Five tools/ gestures selected (knife, spatula/ flipper, tweezers, okay sign, cab hailing) that discriminated performance of stroke patients best from healthy subjects. Based on highest negative Z scores.	Cut-off score: ≥ 2 standard deviations below mean of healthy control group. **[13]**	Z-Scores for most sensitive gestures: Knife (pantomime): -10.0; knife and spatula (object use) = -11.1, -9.92; tweezers (imitation) = -7.50; tweezers and spatula (imitation) = -7.24, -10.09; okay sign and cab hailing (pantomime) = -9.69, -7.24; okay sign and cab hailing (imitation) = -13.49, -18.03 **[13]**	Normative sample: 30 control participants (age and gender matched to patient population) 37 stroke patients, movements performed with ipsilesional limb;, performed all gestures with left and right limb in counterbalanced order.	Presents 5 tools/ gestures out of a sample of 16 items that are best for detecting apraxia. Materials described in: English **[13]**	More comprehensive assessment battery might be required after detecting apraxia. No interrater-reliability reported.

Scoring (videotaped): Five dimension scoring system (action, location, orientation, plane, posture).			Participants originated from Canada **[13]**		
Dementia Apraxia Test (DATE) [14] Test consists of 5 subscales. Imitation of meaningless hand postures and pantomiming of common objects are two subscales for limb apraxia (DATE 1). Imitation of face postures, emblematic buccofacial postures and the repetition of pseudowords are three subscales for buccofacial apraxia (DATE 2). Entails 20 items in total. Time to administer: 10 minutes Scoring (non-videotaped): 3 = target posture achieved smoothly and target-oriented; 2 = halting movement/ spatial corrections; 1 = correct target posutre achieved after verbal cue	Internal consistency: α = 0.84 Interrater reliability (without videotape by 1st examiner; with videotape by 2nd examiner): 0.885 (kappa) Construct validity: r = 0.48 (correlation with KAS [10]) **[14]**	Cut-off score: 45/60 for identifying dementia variants (Sensitivity: 0.91; Specificity: 0.71) Difference score between two subtests (limb apraxia minus buccofacial apraxia) used for differential diagnosis of AD and bvFTD. (Cut-off score: -7; Sensitivity = 0.74; Specificity = 0.93). **[14]** Four dementia subtypes (AD, bvFTD, svPPA, nfPPA) could be correctly discriminated in 64.4% of cases due to praxis-profiles of DATE. **[46]**	Normative data: German sample, 35 healthy subjects (mean age: 67.9). 28 patients with AD (mean age: 71.5); 24 patients with bvFTD (mean age: 64.9); **[14]** *Further study* examined 27 patients with probable AD (age: 61 – 81), 31 patients with probable bvFTD (age: 56 – 72), 21 patients with svPPA (age: 58 – 72), patients with nfPPA (age: 57 – 79) and 34 controls (age: 62 – 74). **[46]**	Items selected based on values to discriminate between AD, bvFTD and healthy subjects first and on psychometric properties afterwards. Same sample was used for the two selection processes. Materials described in English: **[14],** German material available via email to corresponding author	Age difference between AD and bvFTD patients in sample.
Apraxia subtest of Oxford Cognitive Screen (OCS) [15] Test includes imitation of two finger positions and two sequences of hand positions relative to the head. Subjects have to imitate gestures with their better hand. Time to administer (entire battery): 1h Scoring (non-videotaped): 3 = gesture correct after first presentation; 2 = gesture correct after second presentation; 1 = error after second presentation; 0 = no response/ >1 error.	Test-retest reliability: *r* = 0.575 (3 days on average) Convergent validity: *r* = 0.648 (Correlation with BCoS [41] imitation task) Divergent validity: r = 0.042 (Correlation with Barthel index [40]) **[15]**	Incidence of impairments (up to 3 weeks post-stroke): 29.76% LBD; 13.86% RBD; Cut-off: > 5th percentile (score = 8/12); Sensitivity: 72.20%; Specificity: 90.70%. **[15]**	Normative sample: 140 healthy subjects tested at the University of Oxford, UK (age: 36 – 88); 208 acute stroke patients (age: 25 – 96; LBD: 84; RBD: 101; bilateral: 19; unknown: 4). **[15]**	English manual available with clear instructions for application and scoring. Translated and re-normed in several languages: Italian, Cantonese, Danish, Dutch, Putonghua, Russian, Brazilian Portuguese Materials available online: http://www.ocs-test.org/	Includes imitation of meaningless gestures only.

TESTS FOR ASSESSING DCD					
Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2) [16] Test entails four motor area composites (Manual control, manual coordination, body coordination, strength and agility). Two subtests are included in each motor area composite. Time to administer: 40 – 60 minutes (Complete Form); 15 – 20 minutes (Short Form) Scoring (non-videotaped): E.g. number of correct responses/ number of seconds an activity is sustained.	Interrater reliability (non-videotaped): > 0.90 for all subtests apart from Fine Motor Precision (r = 0.86) Test-retest reliability (7 to 42 days): ≥0.80 Internal consistency: ≥0.93 (stratified alpha) Convergent validity: r = 0.80 (Correlation with BOTMP [32]). r = 0.73 (Correlation PDMS-2 [33]) r = 0.74 (Correlation BOT-2 [16] Fine Motor Integration subtest and the TVMS-R [34] Visual-Motor Skills composite) **[16]**	Children in clinical groups (DCD, autism, mental retardation) showed significantly lower scores ($p < .001$) than children in non-clinical comparison group. Clinical group was age-matched with non-clinical group. **[16]**	Normative sample: 1520 US children and youth Normative data reported for 12 age groups from 4 – 21 years. Clinical samples (for discriminative validity): DCD (n = 50); high-functioning autism/Asperger's disorder (n = 45); mild to moderate mental retardation (n = 66). **[16]**	Norms reflect characteristics of US population. Materials described in English: **[16]**	Scoring systems are complex and prone to errors. However, software is available to simplify scoring.
Movement ABC test (M-ABC): [17] Test contains eight different tasks covering aspects of motor coordination that underlie different activities of daily living for children. Tasks fall into three categories: "Manual Dexterity", "Aiming & Catching" & "Balance". Tasks increase in difficulty with age. Time to administer: 20 –30 minutes Scoring (non-videotaped): raw scores depend on "time in seconds", "number of errors" & "number of correct attempts"	Internal consistency: α = .70 - .87 [22] Interrater Reliability (videotaped): .95 – 100 (kappa) [18] Test-retest: (one week): .95 (kappa) [19] Concurrent validity: *r* = .62 [20] (Correlation with Körper Koordinationstest für Kinder (KTK) [21])	Cut-off score: 15th percentile. Scores for different age bands are reported (3 – 6 years; 7 – 10 years; 11 – 16 years). 21 of 34 children with DCD scored below cut-off. 11 of 38 children with nonDCD scored below cut-off [25].	Normative sample: Stratified sample of children from the UK (age: 3 – 16; N = 1172) [17].	More sensitive in identifying children with motor problems than BOT-2 [25]. Materials described in English: **[17]**	Norms might not be generalizable to all cultures (e.g. [23]).
DCD Questionnaire (DCDQ) [26] 15-item questionnaire identifying functional skills in different contextual areas. Asks parents to compare their child´s coordination with children of the same age on a 5-point Likert Scale. Time to administer: 10 - 15 minutes	Internal consistency: α = 0.89 Test-retest reliability (found in Chinese version): .94 [27] Concurrent Validity: significant correlations with M-ABC [17] (r = -.55) and with VMI [42] (r = .42). No correlation with ADHD diagnosis (r = -.11, p = .12). **[26]**	DCD/ suspected DCD group (N = 136) scored significantly lower on the DCDQ'07 than nonDCD group (N = 96) [28]. Age-specific cutoff scores used; sensitivity = 84%; specificity = 71%. **[26]**	Normative sample: Children without DCD (N = 96), with ADHD (N = 35), children with DCD/ suspected DCD (N = 136). Age range: 5.1 – 15.6. Their parents filled out the questionnaire. **[26]**	Cross-cultural adaptations exist (Dutch, Hebrew, Chinese, Portuguese & Greek). Manual of DCDQ'07 available on: https://www.dcdq.ca/	Further standardized tests are needed to complement identification of motor performance deficits in children.

Test of Ideational Praxis (TIP) [29, 30] Assesses ideational skills in children. Child is instructed to show examiner everything he/ she can do with an object out of a series of 6 objects. There is a 5-minute time limit for each item. Time to administer: 30 minutes Scoring (non-videotaped): total number of ways in which the child interacted with objects.	Interrater reliability (non-videotaped): .85 (ICC) [29] Internal consisteny: α = .74**)** [29] Test-retest: (2 weeks) α = .80 (for item "shoestring" only) [30]	Means and standard deviations by age and gender for normative sample are reported. [29] On average, prescoolers (age: 3 – 5, N = 78) interacted in 10.6 different ways with item "shoestring". [30]	Normative sample: 84 healthy children (age: 5 – 8) Children were obtained through a convenience sample (N = 59) and from local (US) school system (N = 25). [29] *Further study* investigating only the item "shoestring" was conducted with 78 children (age: 3 – 5 years). [30]	First objective measurement for assessing ideational abilities in children. Materials described in: English [29]	Scoring is challenging and requires training. [30] Normative sample: Only 10 children per group per (age/year and gender specific). [29]
Abilità Prassiche e della Coordinazione Motoria—2nd Edition (APCM-2) [52, 31] Assesses motor and praxic coordination. Suitable for early identification of motor coordination disorder risk signs in children. Test includes the two sub-scales *motor schemes* (balance and coordination, oculo-motivity, sequencing, and hand and fingers movement), and *adaptive cognitive functions* (dynamic coordination, graphomotor skills, manual skills, symbolic gestures, and constructive praxis abilities). Time to administer: -	Internal consistency: α > .75 for both sub-scales		Normative data: 261 children (age: 3 – 8). Sample consisted of 54% boys and 46% girls.	Six versions adapted for specific age groups available. Materials described in: Italian [52]	Test materials available in Italian only.

General comments.

Psychometric properties: While the reported psychometric properties of the listed tests at large demonstrates satisfying properties, it needs to be pointed out that in most cases the data is based on small samples only. If a differentiation between demographic variables has been considered then even fewer people remain per normative bin. Further, many subtests assessing limb apraxic symptoms evoke ceiling effects in healthy samples (especially tasks with familiar gestures or familiar tool use). Psychometric data and Cut-Off Scores only based on data available from healthy subjects are therefore difficult to interpret. There is a clear need for larger sample sizes, including patients, and for extending methodological evaluations.

A few additional notes on the evaluation of apraxic behavior based on clinical and teaching experience: Typically, experience is needed for evaluating apraxic performance online while the patient is being tested. Before evaluating a patient's apraxic performance, testing personnel should be familiar with the scoring system and criteria of the chosen test. Complex behaviors and settings, such as naturalistic actions including tools and objects, draw upon the clinician's / experimenter's capacities. Frequently, a quick change of items between trials is needed, instructions need to be given and the patient needs to be monitored at the same time. Scoring of apraxic behavior oftentimes requires looking at performance several times. Therefore, especially in the beginning, the use of video-tapes can be helpful, allowing to postpone and review the evaluation of the patient's performance. It is highly recommended to study the provided training material

(e.g. manuals, videos). Please also note, that each evaluation represents only a snapshot of the patient's behaviour. Patients may even demonstrate inconsistent behavior per item. Items performed erroneously in one run may be fine in the next session and vice versa. Problems demonstrated in brief screenings should be followed up by more elaborate testing and observation.

Note. HC = Healthy controls; TBI = Traumatic brain injury; IMA = Ideomotor apraxia; CA = conceptual apraxia; LKA = Limb kinetic apraxia; AD = Alzheimer's dementia; bvFTD = behavioral variant frontotemporal dementia, nfPPA = nonfluent primary progressive aphasia; svPPA = semantic variant primary progressive aphasia; DCD = developmental coordination disorder; IADL = "instrumental" ADL; α = Cronbach's alpha; SD = standard deviation

1. Power, E., Code, C., Croot, K., Sheard, C., & Gonzalez Rothi, L. J. (2010). Florida Apraxia Battery–Extended and Revised Sydney (FABERS): Design, description, and a healthy control sample. *Journal of clinical and experimental neuropsychology*, *32*(1), 1-18. https://doi.org/10.1080/13803390902791646
2. Vanbellingen, T., Kersten, B., Van Hemelrijk, B., Van de Winckel, A., Bertschi, M., Müri, R., ... & Bohlhalter, S. (2010). Comprehensive assessment of gesture production: a new test of upper limb apraxia (TULIA). *European Journal of Neurology*, *17*(1), 59-66. https://doi.org/10.1111/j.1468-1331.2009.02741.x
3. De Renzi, E., Motti, F., & Nichelli, P. (1980). Imitating gestures: a quantitative approach to ideomotor apraxia. *Archives of neurology*, 37(1), 6-10. 10.1001/archneur.1980.00500500036003
4. Kertesz A, Ferro JM. (1984). Lesion size and location in ideomotor apraxia. *Brain*, 3: 921–933. https://doi.org/10.1093/brain/107.3.921
5. Bartolo, A., Cubelli, R., & Sala, S. D. (2008). Cognitive approach to the assessment of limb apraxia. *The Clinical Neuropsychologist*, *22*(1), 27-45.
6. Schwartz, M. F., Segal, M., Veramonti, T., Ferraro, M., & Buxbaum, L. J. (2002). The Naturalistic Action Test: A standardised assessment for everyday action impairment. *Neuropsychological rehabilitation*, *12*(4), 311-339. https://doi.org/10.1080/13854040601139310
7. Randerath, J., Buchmann, I., Liepert, J., & Büsching, I. (2017). Diagnostic Instrument for Limb Apraxia: Short Version (DILA-S). http://nbn-resolving.de/urn:nbn:de:bsz:352-2-1uq46vs8q2f0i8
8. Buchmann, I., Dangel, M., Finkel, L., Jung, R., Makhkamova, I., Binder, A., ... & Randerath, J. (2020). Limb apraxia profiles in different clinical samples. *The Clinical Neuropsychologist*, *34*(1), 217-242. https://doi.org/10.1080/13854046.2019.1585575
9. Vanbellingen, T., Kersten, B., Van de Winckel, A., Bellion, M., Baronti, F., Müri, R., & Bohlhalter, S. (2011). A new bedside test of gestures in stroke: the apraxia screen of TULIA (AST). *Journal of Neurology, Neurosurgery & Psychiatry*, *82*(4), 389-392. http://dx.doi.org/10.1136/jnnp.2010.213371
10. Weiss-Blankenhorn, P. H., & Eschenbeck, P. (2014). Kölner Apraxie-Screening–Praktisch bei Apraxie. *Ergopraxis*, *7*(07/08), 38-39. 10.1055/s-0034-1386567
11. Leiguarda, R., Clarens, F., Amengual, A., Drucaroff, L., & Hallett, M. (2014). Short apraxia screening test. *Journal of clinical and experimental neuropsychology*, *36*(8), 867-874. https://doi.org/10.1080/13803395.2014.951315
12. Tessari, A., Toraldo, A., Lunardelli, A., Zadini, A., & Rumiati, R. I. (2015). STIMA: a short screening test for ideo-motor apraxia, selective for action meaning and bodily district. *Neurological Sciences*, *36*(6), 977-984. https://doi.org/10.1007/s10072-015-2203-4
13. Almeida, Q. J., Black, S. E., & Roy, E. A. (2002). Screening for apraxia: a short assessment for stroke patients. *Brain and cognition*, *48*(2-3), 253–258. https://doi.org/10.1006/brcg.2001.1356
14. Johnen, A., Frommeyer, J., Modes, F., Wiendl, H., Duning, T., & Lohmann, H. (2016). Dementia Apraxia Test (DATE): a brief tool to differentiate behavioral variant frontotemporal dementia from Alzheimer's dementia based on apraxia profiles. *Journal of Alzheimer's Disease*, *49*(3), 593-605. 10.3233/JAD-150447
15. Demeyere, N., Riddoch, M. J., Slavkova, E. D., Bickerton, W. L., & Humphreys, G. W. (2015). The Oxford Cognitive Screen (OCS): validation of a stroke-specific short cognitive screening tool. *Psychological assessment*, *27*(3), 883. 10.1037/pas0000082
16. Deitz, J. C., Kartin, D., & Kopp, K. (2007). Review of the Bruininks-Oseretsky test of motor proficiency, (BOT-2). *Physical & occupational therapy in pediatrics*, *27*(4), 87-102. https://doi.org/10.1080/J006v27n04_06
17. Schulz, J., Henderson, S. E., Sugden, D. A., & Barnett, A. L. (2011). Structural validity of the Movement ABC-2 test: Factor structure comparisons across three age groups. *Research in developmental disabilities*, *32*(4), 1361-1369. https://doi.org/10.1016/j.ridd.2011.01.032
18. Smits-Engelsman, B. C., Fiers, M. J., Henderson, S. E., & Henderson, L. (2008). Interrater reliability of the movement assessment battery for children. *Physical therapy*, *88*(2), 286-294. https://doi.org/10.2522/ptj.20070068
19. Croce, R. V., Horvat, M., & McCarthy, E. (2001). Reliability and concurrent validity of the movement assessment battery for children. *Perceptual and motor skills*, *93*(1), 275-280. https://doi.org/10.2466/pms.2001.93.1.275
20. Smits-Engelsman, B. C., Henderson, S. E., & Michels, C. G. (1998). The assessment of children with Developmental Coordination Disorders in the Netherlands: The relationship between the Movement Assessment Battery for Children and the Körperkoordinations Test für Kinder. *Human Movement Science*, *17*(4-5), 699-709. https://doi.org/10.1016/S0167-9457(98)00019-0
21. Kiphard, E. J., & Schilling, F. (2007). *Körperkoordinationstest für Kinder: KTK*. Beltz-Test. https://doi.org/10.1016/S0167-9457(98)00019-0
22. Smits-Engelsman, B. C., Niemeijer, A. S., & van Waelvelde, H. (2011). Is the Movement Assessment Battery for Children-a reliable instrument to measure motor performance in 3 year old children? *Research in developmental disabilities*, *32*(4), 1370-1377. https://doi.org/10.1016/j.ridd.2011.01.031
23. Miyahara, M., Tsujii, M., Hanai, T., Jongmans, M., Barnett, A., Henderson, S. E., ... & Kageyama, H. (1998). The Movement Assessment Battery for Children: A preliminary investigation of its usefulness in Japan. *Human Movement Science*, *17*(4-5), 679-697. https://doi.org/10.1016/S0167-9457(98)00018-9
24. Bruininks, R. H. (1978). *Bruininks-Oseretsky test of motor proficiency*. Circle Pines, MN: American Guidance Service.
25. Crawford, S. G., Wilson, B. N., & Dewey, D. (2001). Identifying developmental coordination disorder: consistency between tests. *Physical & occupational therapy in pediatrics*, *20*(2-3), 29-50. https://doi.org/10.1080/J006v20n02_03
26. Wilson, B. N., Kaplan, B. J., Crawford, S. G., & Roberts, G. (2007). The developmental coordination disorder questionnaire 2007 (DCDQ'07). *Administrative manual for the DCDQ107 with psychometric properties*.

27. Tseng, M. H., Fu, C. P., Wilson, B. N., & Hu, F. C. (2010). Psychometric properties of a Chinese version of the Developmental Coordination Disorder Questionnaire in community-based children. *Research in Developmental Disabilities*, *31*(1), 33-45. https://doi.org/10.1016/j.ridd.2009.07.018
28. Wilson, B. N., Crawford, S. G., Green, D., Roberts, G., Aylott, A., & Kaplan, B. J. (2009). Psychometric properties of the revised developmental coordination disorder questionnaire. *Physical & occupational therapy in pediatrics*, *29*(2), 182-202. https://doi.org/10.1080/01942630902784761
29. May-Benson, T. A., & Cermak, S. A. (2007). Development of an assessment for ideational praxis. *American Journal of Occupational Therapy*, *61*(2), 148-153. https://doi.org/10.5014/ajot.61.2.148
30. Lane, S. J., Ivey, C. K., & May-Benson, T. A. (2014). Test of Ideational Praxis (TIP): Preliminary findings and interrater and test–retest reliability with preschoolers. *American Journal of Occupational Therapy*, *68*(5), 555-561. https://doi.org/10.5014/ajot.2014.012542
31. Minoliti, R., Crepaldi, M., & Antonietti, A. (2020). Identifying developmental motor difficulties: a review of tests to assess motor coordination in children. *Journal of Functional Morphology and Kinesiology*, *5*(1), 16. https://doi.org/10.3390/jfmk5010016
32. Bruininks-Oseretsky, R. (1978). Test of motor proficiency: Examiner's manual. *Circle Pines, MN: American Guidance Service*.
33. Folio, M. R., & Fewell, R. R. (2000). PDMS-2 Peabody developmental motor scales second edition. *Austin: PRO-ED Inc*.
34. Gardner, M. F. (1995). *Test of visual-motor skills-revised*. Psychological and Educational Publications.
35. Rothi, L. J. G., Raymer, A. M., & Heilman, K. M. (1997). Limb praxis assessment. *Apraxia: The neuropsychology of action*, 61-73.
36. Kidd, D., Stewart, G., Baldry, J., Johnson, J., Rossiter, D., Petruckevitch, A., & Thompson, A. J. (1995). The Functional Independence Measure: a comparative validity and reliability study. *Disability and rehabilitation*, *17*(1), 10-14. https://doi.org/10.3109/09638289509166622
37. Wilson, B. A., Cockburn, J., & Halligan, P. W. (2004). *Behavioural inattention test*. École des sciences de la réadaptation, Sciences de la santé, Université d'Ottawa.
38. McDowell, S., Whyte, J., & D'Esposito, M. (1997). Working memory impairments in traumatic brain injury: evidence from a dual-task paradigm. *Neuropsychologia*, *35*(10), 1341-1353. https://doi.org/10.1016/S0028-3932(97)00082-1
39. Rothi, L. G., Raymer, A. M., Ochipa, C., Maher, L. M., Greenwald, M. L., & Heilman, K. M. (1992). Florida apraxia battery, experimental edition. *Gainesville, FL: Univ of Florida College of Medicine*.
40. Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: the Barthel Index: a simple index of independence useful in scoring improvement in the rehabilitation of the chronically ill. *Maryland state medical journal*.
41. Humphreys, G. W., Bickerton, W. L., Samson, D., & Riddoch, M. J. (2012). *BCoS cognitive screen*.
42. Beery, K. E. (2004). Beery VMI: The Beery-Buktenica developmental test of visual-motor integration. *Minneapolis, MN: Pearson*.
43. Dovern, A., Fink, G. R., & Weiss, P. H. (2012). Diagnosis and treatment of upper limb apraxia. *Journal of neurology*, *259*(7), 1269-1283. https://doi.org/10.1007/s00415-011-6336-y
44. Vanbellingen, T., Lungu, C., Lopez, G., Baronti, F., Müri, R., Hallett, M., & Bohlhalter, S. (2012). Short and valid assessment of apraxia in Parkinson's disease. *Parkinsonism & related disorders*, *18*(4), 348-350. https://doi.org/10.1016/j.parkreldis.2011.11.023
45. Goetz, C. G., Tilley, B. C., Shaftman, S. R., Stebbins, G. T., Fahn, S., Martinez-Martin, P., ... & LaPelle, N. (2008). Movement Disorder Society-sponsored revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): scale presentation and clinimetric testing results. *Movement disorders: official journal of the Movement Disorder Society*, *23*(15), 2129-2170. https://doi.org/10.1002/mds.22340
46. Johnen, A., Reul, S., Wiendl, H., Meuth, S. G., & Duning, T. (2018). Apraxia profiles—A single cognitive marker to discriminate all variants of frontotemporal lobar degeneration and Alzheimer's disease. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, *10*, 363-371. https://doi.org/10.1016/j.dadm.2018.04.002
47. Goldenberg, G. (1995). Imitating gestures and manipulating a mannikin—the representation of the human body in ideomotor apraxia. *Neuropsychologia*, *33*(1), 63-72. https://doi.org/10.1016/0028-3932(94)00104-W
48. De Renzi, E., Pieczuro, A., & Vignolo, L. A. (1968). Ideational apraxia: a quantitative study. *Neuropsychologia*, *6*(1), 41-52. https://doi.org/10.1016/0028-3932(68)90037-7
49. Raven, J. C. (1941). Standardization of progressive matrices, 1938. *British Journal of Medical Psychology*, *19*(1), 137-150. https://doi.org/10.1111/j.2044-8341.1941.tb00316.x
50. De Renzi, E., & Faglioni, P. (1965). The comparative efficiency of intelligence and vigilance tests in detecting hemispheric cerebral damage. *Cortex*, *1*(4), 410-433. https://doi.org/10.1016/S0010-9452(65)80003-X
51. Vanbellingen, T. (2013). Assessment: Apraxia Screen of TULIA–Gesten testen. *ergopraxis*, *6*(09), 22-23. 10.1055/s-0033-1356909
52. Sabbadini, L., Tsafrir, Y., & Iurato, E. (2005). Protocollo APCM: valutazione delle Abilità Prassiche e della Coordinazione Motoria per fasce di età. In *Protocollo per la valutazione delle Abilità Prassiche e della Coordinazione Motoria APCM* (pp. 23-56). Springer, Milano.