
**Exemplary Tests**

**CLASSICAL TEST BATTERIES (with focus on gestures)**

<table>
<thead>
<tr>
<th>Test of Upper Limb Apraxia (TULIA)</th>
<th>Psychometrics</th>
<th>Clinical Sensitivity</th>
<th>Normative data</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
</table>
| TULIA                              | 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  

**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 24h) [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:** 80 healthy subjects originating from Switzerland and Germany: 43 – 93. 84 subjects with left brain damage (age: 30 – 84), 49 subjects with right brain damage (age: 33 – 85), [2]

- **Short time administration; high reliability and validity; detailed scoring system.**
- **Materials published in English:** [2]

- **Imitation of nonsymbolic gestures in part less reliable.** [2]
| Test of Bartolo et al. [5] | 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 (15.9/5/416) and visual (15.5/4/16). Production of Pantomimes: Verbal (15.6/4/16), visual (15.7/4/16) and tactile (15.6/4/16). Recognition: intransitive gestures (39.2/4/16) and pantomimes (39.9/4/40). Remaining tasks performed at ceiling: transitive gestures | Normative data: 60 British subjects (age: 17–41); healthy, right-handed. Materials described in English. [5] | 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. |
| 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. 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: 40 control patients without brain injury tested in Italy (mean age: 52.6); 40 RBD (mean age: 57.7); 40 LBD (mean age: 65.9). Patients carried out task with ipsilateral side of lesion. Half of control 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. | Time to administer: 15 minutes | Scoring (non-videotaped): 3, 2, 1 or 0 points (based on number of trials needed) |
| De Renzi “idiotational” 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 “idiotational apraxia” subset and an “ideo-motor” apraxia subset. | External validity: Rho = .410 (Cov. wih iitemo-motor subtest) Rho = .392 (Cov. with Raven’s progressive matrices [49]) Rho = .263 (Cov. with visual reaction times [50]) Rho = .579 (Cov. 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. LMA subtest: Cutoff score: > 17 points (total score: 20). No control/ RBD scored under cutoff. 45 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] | Time to administer: - | Test includes gestures that are culture-specific to Italy. Ceiling effects especially for IA subset. |
### 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 item 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 (NTI) 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-video-taped; 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 = .590;
  - 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].

**Normative sample:**
- German sample, 82 right-handed healthy subjects tested uni-manually (age: 11 – 80);
- Psychometrics based on: 13 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.mooc.uni-konstanz.de/publikationen/assessments/](https://www.mooc.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-video-taped):** Each item scored for accomplishment and errors.

**Interrater reliability (without videotape by 1st examiner, with videotape by 2nd examiner):**
- 0.98 (kappa)

**Internal consistency:**
- τ = 0.753 (for patient sample);
- τ = 0.793 (for entire sample).

**Concurrent validity:**
- Significant correlations 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:**
- 18 non-neurological control subjects tested in the USA (age: 18 – 80);

**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.

<table>
<thead>
<tr>
<th>Limb Apraxia Screenings:</th>
<th>LIMB APRAXIA SCREENINGS:</th>
</tr>
</thead>
</table>
| **Apraxia Screen of TULIA (AST)** [9, 44] | **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 par. III, [15] (measuring parkinsonism symptoms; \( r = 0.17, p = 0.1 \)) [44]  
**Convergent validity:** association with MDS-UPDRS par. II/III [45] (measuring ADL; \( r = 0.35 \)) and ;loehn & Yah: 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:** 17 RBD and 26 LBD stroke patients (average age: 63.4). All subjects were right-handed. [9]  
**Short time to administer:** 3 minutes  
**Materials described in:** English [9] and German: [51]  
**Small sample size used for validation:** |
| **Kölner Apraxia Screening (KAS)** [10, 43] | **Internal consistency:** \( \alpha = 0.968 \)  
**Homogeneity:** 0.604  
**Inter-rater reliability (videotaped):** \( \rho = 0.907 \)  
**Construct validity:** 653 and .565 with two subscales of an imitation test of Goldberg [47]  
**Cut-off:** \( \leq 76 \) (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:** 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:** [43]  
**Scoring (non-videotaped):** 4 – 0 points for each item (depending on item-specific scoring criteria)  
**Time to administer:** 10 Minutes  
**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.**
<table>
<thead>
<tr>
<th>Test Name</th>
<th>Description</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Short test for ideo-motor apraxia (STIMA)</strong> [12]</td>
<td>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.</td>
<td>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]</td>
</tr>
</tbody>
</table>
**Scoring (videotaped):** Five dimension scoring system (action, location, orientation, plane, posture).

**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). Examines 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 posture achieved after verbal cue

**Internal consistency:** $\alpha = 0.84$

**Interrater reliability (without videotape by 1st examiner, with videotape by 2nd examiner):** 0.885 (kappa)

**Construct validity:** $r = 0.48$ (correlation with KAS [10])

**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, nPPA) 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); 38 patients with AD (mean age: 71.5); 24 patients with bvFTD (mean age: 64.9). [14]


**Participants originated from Canada [13]**

**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].

**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; ( = no response/ > 1 error.

**Test-retest reliability:** $r = 0.575$ (3 days on average)

**Convergent validity:** $r = 0.648$

(Conrelation with BCs [41] imitation task)

**Divergent validity:** $r = 0.042$

(Conrelation with Barthel index [46])

**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:**
40 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.ocstest.org/](http://www.ocstest.org/)

**Includes imitation of meaningless gestures only.**
<table>
<thead>
<tr>
<th>TESTS FOR ASSESSING DCD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bruininks-Oseretsky Test of Motor Proficiency, Second Edition (BOT-2)</strong> [16]</td>
</tr>
<tr>
<td>Test entails four motor area composites (manual control, manual coordination, body coordination, strength and agility). Two subtests are included in each motor area composite.</td>
</tr>
<tr>
<td><strong>Time to administer:</strong> 40 – 60 minutes (Complete Form); 15 – 20 minutes (Short Form)</td>
</tr>
<tr>
<td>Scoring (non-videotaped): E.g. number of correct responses/number of seconds an activity is sustained.</td>
</tr>
<tr>
<td>Intercoder reliability (non-videotaped): &gt; 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 BCITMP [32]; r = 0.73 (Correlation with PDMS-2 [33]; r = 0.74 (Correlation with VMS-R [34]; Visual-Motor Skills composite) [16] Children in clinical groups (DCD, autism, mental retardation) showed significantly lower scores (p &lt; .001) than children in non-clinical comparison group. Clinical group was age-matched with non-clinical group. [16] Normative sample: 520 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 = 55); 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.</td>
</tr>
<tr>
<td><strong>Movement ABC test (M-ABC):</strong> [17]</td>
</tr>
<tr>
<td>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 &amp; Catching” &amp; “Balance”. Tasks increase in difficulty with age.</td>
</tr>
<tr>
<td><strong>Time to administer:</strong> 20 – 30 minutes</td>
</tr>
<tr>
<td>Scoring (non-videotaped): raw scores depend on “time in seconds”, “number of errors” &amp; “number of correct attempts”</td>
</tr>
<tr>
<td>Internal consistency: α = 0.70 - .87 [22] Intercoder reliability (videotaped): .95 (Kappa) [18] Test-retest (one week): 0.95 (Kappa) [19] Concurrent validity: r = .62 [20] (Correlation with Körner 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 [25]. Materials described in English [25]. More sensitive in identifying children with motor problems than BOT 2 (25). Norms might not be generalizable to all cultures (e.g. [23]).</td>
</tr>
<tr>
<td><strong>DCD Questionnaire (DCDQ):</strong> [26]</td>
</tr>
<tr>
<td>15-item questionnaire identifying functional skills in different contextual areas. Ask parents to compare their child’s coordination with children of the same age on a 5-point Likert Scale.</td>
</tr>
<tr>
<td><strong>Time to administer:</strong> 10 - 15 minutes</td>
</tr>
</tbody>
</table>
| Test of Ideational Praxis (TIP) [29, 30] | Interrater reliability (non-videotaped): 85 (ICC) [29]  
Internal consistency: \( \alpha = .74 \) [29]  
Test-retest, (2 weeks) \( \alpha = .80 \) (for item “shoestring” only) [30] | Means and standard deviations by age and gender for normative sample are reported. [29]  
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/sex and gender specific. [29] |
|-----------------------------|---------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------------------------------------------|
| **Abilità Prassiche e della Coordinazione Motoria—2nd Edition (APCM-2) [5, 31]** | Internal consistency: \( \alpha > .75 \) 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. | |

**Time to administer:** 30 minutes  
**Scoring (non-videotaped):** total number of ways in which the child interacted with objects.

**Abilità Prassiche e della Coordinazione Motoria—2nd Edition (APCM-2)**
Test includes the two sub-scales *motor schemes* (balance and coordination, oculo-motility, 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:**  
**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 differentiation between demographic variables has been considered then even fewer people remain per normative bin. Further, many subtests assessing limb apraxia 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; nPPA = nonfluent primary progressive aphasia; svPPA = semantic variant primary progressive aphasia; DCD = developmental coordination disorder; IADL = “instrumental” ADL; α = Cronbach’s alpha; SD = standard deviation*


