# A new bedside test of gestures in stroke: the apraxia screen of TULIA (AST)

T Vanbellingen,[1] B Kersten,[2] A Van de Winckel,[3] M Bellion,[1] F Baronti,[1] R Müri,[4] S Bohlhalter[4]

► Appendix 1 and 2 are published online only. To view these files please visit the journal online (http://jnnp.bmj.com).

[1]Klinik Bethesda, Neurorehabilitation Centre, Tschugg, Switzerland
[2]Department of Psychology, University of Bern, Bern, Switzerland
[3]Department of Rehabilitation Sciences, KU Leuven, Belgium
[4]Division of Cognitive and Restorative Neurology, Department of Neurology, University Hospital Bern, Bern, Switzerland

**Correspondence to**
Dr Stephan Bohlhalter, Division of Cognitive and Restorative Neurology, Department of Neurology, Inselspital, Bern University Hospital, Bern, Switzerland;
stephan.bohlhalter@insel.ch

Received 25 March 2010
Revised 5 July 2010
Accepted 21 July 2010
Published Online First
9 October 2010

## ABSTRACT

**Background** Apraxia in patients with stroke may be overlooked, as clumsiness and deficient gestural communication are often attributed to frequently coexisting sensorimotor deficits and aphasia. Early and reliable detection of apraxia by a bedside test is relevant for functional outcome in patients with stroke. The present study was aimed at constructing a new bedside screening test for apraxia, called the Apraxia Screen of TULIA (AST), based on the comprehensive standardised Test for Upper-Limb Apraxia (TULIA).

**Methods** First, an item-reduction analysis of the TULIA (48 gestures) was performed, based on the methods of classical test theory and on a larger sample of patients with stroke (n=133) and matched healthy controls (n=50). Stepwise elimination of items resulted in a set of 12 items, demonstrating high internal consistency (Cronbach alpha=0.92). The six-point scoring method of the TULIA was dichotomised to the score levels pass and fail. In the second part of this study the validity of the AST was assessed prospectively in a new cohort of patients with stroke (n=31) by using the Pearson correlation analysis and binary classification display with the TULIA.

**Results and discussion** Validation of the 12-item AST with the TULIA showed a remarkable diagnostic reliability with high specificity, sensitivity and positive predictive value, for the presence and severity of apraxia. The AST is shown to be a reliable and valid bedside test in patients with stroke, allowing a straightforward assessment of apraxia within a few minutes.

Limb apraxia is a common higher-order motor dysfunction, mainly associated with left-hemisphere brain damage.[1–3] It is characterised by an inability to perform skilled and/or learnt limb movements correctly, not accounted for by an elementary motor and sensory disorder, or cognitive deficit.[4] Several studies have shown that apraxia impedes activities of daily living (ADL) and can lead to a persistent disability in patients with stroke.[5–9] The functional impact of apraxia may be less obvious when a stroke patient performs a familiar everyday activity inappropriately (eg, the use of a toothbrush to eat spaghetti) only on few occasions.[10 11] Furthermore, patients may not appear to have obvious problems with routine (basic) daily activities, but their use of tools and objects is inaccurate, inefficient or even hazardous.[9 12] In addition, clumsiness, particularly in bimanual tasks, can be viewed as a reflection of the frequently coexisting hemiparesis, thus masking the diagnosis of an apraxia. Finally, many left-hemispheric patients with stroke are aphasic; the presence of an apraxia can restrict the use of gestures to compensate for their expressive language deficits.[13 14] A bedside screening test that allows for a quick and reliable diagnosis of apraxia is therefore of particular significance.

To our knowledge, only one published study has introduced a screening test for apraxia,[15] but the validation of this study is incomplete. A major drawback of this screening tool is that the scoring method is rather complex for bedside use. The comprehensive standardised test for upper-limb apraxia (TULIA),[1] which we recently developed, was also not conceived to be used for a bedside evaluation of apraxia but provided the basis for the development of a screening version. Hence, the aim of the present study was twofold: first to construct a new bedside test for limb apraxia in patients with stroke by performing an item-reduction analysis of the TULIA, in the following referred to as the Apraxia Screen of TULIA (AST); second to assess prospectively the diagnostic accuracy of the AST test in a cohort of patients with stroke based on TULIA scores.

## SUBJECTS AND METHODS

### Development of AST

The AST was developed in two phases. In a first phase, an item reduction in the TULIA was performed, based on the original sample (n=183)[1]; in the second phase, the validity of the AST was assessed, based on a new sample (n=31). The TULIA consists of 48 items, with an equal number of items testing for imitation and pantomime. In both of these task domains, the respective set of 24 items comprises eight meaningless gestures, eight intransitive (communicative) and eight transitive (tool-related) gestures. Based on separate cut-off scores and a detailed six-point scoring method (see online appendix 1), the TULIA can identify the pattern as well as the severity of an apraxic impairment. Statistical analyses were performed using PASW for Windows (Version 18.0.0; SPSS).

The item-reduction analysis was based on classical test theory.[16] Accordingly, items with less inter-rater reliability, floor and ceiling effects, less internal consistency and lower content validity (based on expert opinion: TV and SB) were excluded stepwise.

The six-point scoring system of the TULIA was simplified to a dichotomous pass or fail as follows. Score levels 5 (normal movement), 4 and 3 of the TULIA were combined in AST to the score 1 (=pass) (see online appendix 1). The decision to consider the TULIA score levels 4 and 3 as 'normal

movement' was based on the fact that these error types may be difficult to discern by non-specialist examiners. They either did not affect trajectory of gestures, that is, overall movement shape (including the plane relative to a tool or own body) or were subtle and corrected during ongoing movement. However, if errors affecting trajectory persisted, even though being only mild (score=2), gesture performance was considered deviant. Thus, the score levels 2, 1 and 0 of TULIA were pooled to the score 0 (=fail), a score that includes major spatial or semantic errors. Examples of apraxic errors observed in patients with stroke have been replicated by a model for different items and score levels; the videos are available at the author's website (http://www.tulia.ch).

Finally, the cut-off levels of AST were determined in the original sample (n=133) by comparing the corresponding 12 items from TULIA recoded to pass and fail with the full version. Accordingly, using cut-off levels of 9 and 5, high specificity (93%) and sensitivity (88%) for mild and severe apraxia could be estimated.

For the validation of the AST test, patients with stroke were examined with both the AST and the TULIA. The diagnostic reliability of the AST was calculated using the Pearson correlation analysis, with the TULIA score and the AST score as dependent variables. Furthermore, a binary classification display of the AST and TULIA was used to demonstrate sensitivity, specificity and predictive values of the AST.

### Subjects

The item-reduction analysis performed in phase I was based on the data we obtained from a sample of 133 patients (84 with a left and 49 with a right-hemispheric stroke) and 50 healthy controls. Details of their clinical characteristics were reported elsewhere.[1] Demographic and clinical data of the new patient sample participating in phase II are summarised in table 1. The patients were selected prospectively, the selection criterion being that they had sustained a single radiologically confirmed stroke (CT or MRI) and did not suffer from any brain disorder; in particular, showed no signs of dementia (according to DSM IV criteria) or Parkinsonism, or suffered from orthopaedic disorders of the upper non-hemiplegic limb that could interfere with the TULIA or AST task. Furthermore, patients with impaired consciousness or an inability to understand basic task instructions were also excluded. All patients in phase II were right-handed, as assessed by the Edinburgh Handedness Inventory.[17] Written informed consent for participation in the study was obtained from all patients according to the Declaration of Helsinki, 1975. The study was approved by the local ethical committee.

**Table 1** Characteristics of patients (phase II)

| **Age (years)** | **63.4 (13.7)*** |
|---|---|
| Gender | |
| Women | 10 |
| Men | 21 |
| Type of stroke | |
| Infarction | 26 |
| Haemorrhage | 5 |
| Hemisphere of stroke | |
| Right | 5 |
| Left | 26 |
| Phase of stroke | |
| Acute (<4 weeks) | 15 |
| Chronic (>4 weeks) | 16 |
| Aphasia (token test, screening version) | |
| Severe | 8 |
| Mild | 4 |
| Duration (poststroke to initial testing; days) | |
| Acute | 19.0 (5.3)* |
| Chronic | 57.3 (25.5)* |
| Total | 36.9 (25.7)* |

*Mean (SD).

### Procedure

All patients were examined by two investigators (TV and SB) with the TULIA test[1] for apraxia, the Bell test[18] for neglect and the 10 item screening version of the TOKEN test[19] for aphasia. Performances of the TULIA were videotaped for later scoring. Within 24 h of examining the patients with these tests, the AST was administered and immediately scored by the same investigators. Scoring of the TULIA was done by experienced raters (FA and CH; see acknowledgements) who were blinded for the results of the AST.

To perform both the TULIA and the AST test subjects were either seated in front of the examiner or sat in an upright position in bed. In both of these tests the patients executed the movements with their non-paretic upper limb. The imitation tasks required the patients to reproduce the movement that the examiner had just demonstrated in a mirrored fashion. In the pantomime tasks the patients were given oral instructions as to the gesture they were to perform (see online appendix 2). The administration of the AST required on average 3 min after the patient was properly seated and had received the necessary instructions. The order of items followed the list in online appendix 2.

## RESULTS

The results of the stepwise item-reduction analysis are summarised in a flow chart (figure 1). Initially, four of the 48 items of the TULIA with lower inter-rater reliability (weighted kappa <0.60) were eliminated. Second, 12 items with a ceiling effect (mean score >4) were excluded. Third, an iterative elimination of 15 items with lower corrected item-to-total correlation coefficients ($r_{it}$<0.60) was carried out. The remaining 17 items had a high reliability coefficient (Cronbach alpha=0.91) for the temporary total scale. Fourth, the five items with lower content validity were eliminated. We were thus left with the 12 items that constitute the AST, seven items in the imitation and five in the pantomime domain. The item-reduction analysis resulted in the elimination of 15 meaningless, 13 intransitive and eight transitive gestures from the set of 48 items that constitute the TULIA. Thus, the AST consists of one meaningless gesture, three intransitive and eight transitive gestures. Internal consistency for the new scale AST (k=12) remained high (Cronbach alpha=0.92).

According to the dichotomous scoring system (1=pass, 0=fail) the maximum score for AST was 12. Furthermore, the cut-off scores were set at 9 and 5 for mild and severe apraxia, respectively (for rationale, see Subjects and methods section). On the basis of AST test scores, 19 patients (61%) were classified apraxic, and 12 patients (39%) non-apraxic. In the group with apraxia, 12 patients (39%) were mildly, and seven patients (23%) severely apraxic. All patients with severe apraxia had sustained a left-hemispheric stroke, whereas the group of mild apaxia also included two patients with a right-hemispheric stroke. As summarised in table 2, mean total scores and domain subscores of AST and TULIA were significantly lower in apraxic than in non-apraxic patients.

Further validation demonstrated a highly significant correlation between AST and TULIA scores (Pearson correlation

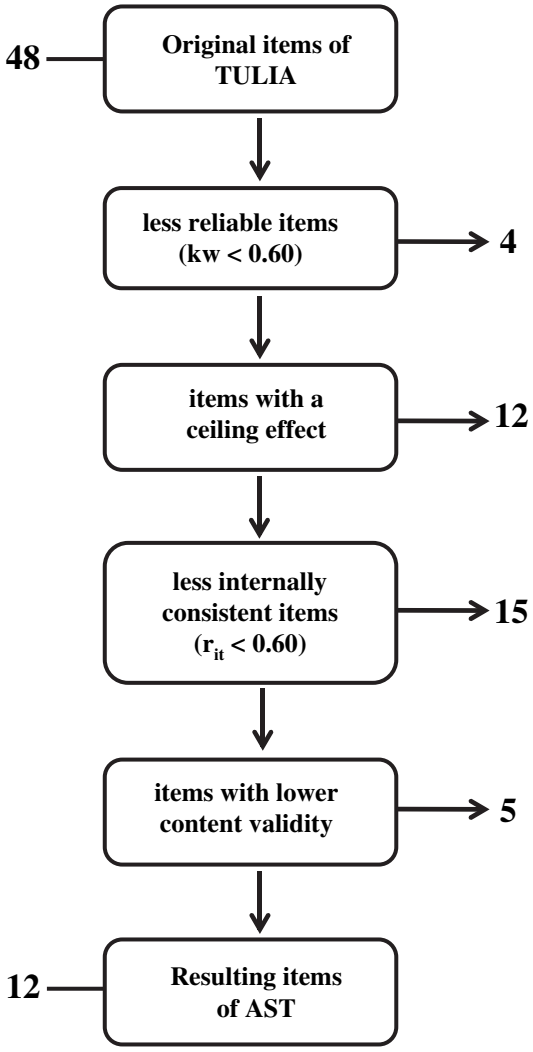


**Figure 1** Flow chart showing the individual steps of the item-reduction analysis with consecutive exclusions (arrows) of items from the Test for Upper-Limb Apraxia (TULIA), finally resulting in the 12 items of Apraxia Screen of TULIA (AST).

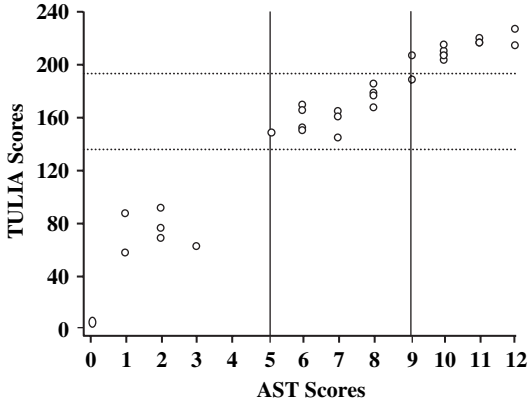


**Figure 2** Scatter plot (with individual scores of 31 patients for both Test for Upper-Limb Apraxia (TULIA) and Apraxia Screen of TULIA (AST)) showing a high correlation between TULIA and AST. Solid lines indicate cut-off scores of AST (9 for mild apraxia and 5 for severe apraxia). Dashed lines indicate cut-off scores of TULIA (194 for mild apraxia and 130 for severe apraxia). Note the almost perfect overlap in non-apraxic, mild and severely apraxic clusters.

coefficient r=0.96, p<0.001) and a largely overlapping classification of patients in a non-apraxic, mildly and severely apraxic (figure 2). The high correlation of the AST with the same 12 items selected from TULIA (r=0.95, p<0.001), recoded to pass and fail, indicated a good test–retest reliability of AST.

More specifically, a binary classification display (table 3) showed that AST's diagnostic accuracy was almost perfect: the specificity of the AST was 100% (11 true negatives), whereas the sensitivity of the AST was slightly lower (19 true positives, 95%), due to the fact that one patient was classified as mildly apraxic by the TULIA but as non-apraxic by the AST. The positive predictive value of the AST was 100%, and the negative predictive value was 92%.

As may be expected, within the group of left-hemispheric patients with stroke (n=26) the aphasic patients (n=12) had significantly lower scores in the AST than the non-aphasic patients (2.9±1.9 and 9.7±0.6, unpaired t test, p<0.01). To assess the potential confound of apraxia by aphasia or neglect, Pearson correlations were calculated between the AST, the Token test and the Bell test in the apraxic subpopulation (n=19). Total AST scores correlated significantly with the Token test (r=0.81, p<0.001). Further analysis showed that the association was particularly strong for the non-verbal imitation domain (r=0.89, p<0.001) and did not reach significance for verbally mediated pantomime domain (r=0.42, p=0.075). There was no relationship between the Bell test scores and the AST total scores and subscores (all p>0.05).

## DISCUSSION

The present study was aimed at constructing a validated bedside test for apraxia in patients with stroke, referred to as the AST. The construction of the AST was based on an item-reduction analysis of a comprehensive test for gestures, called TULIA. For this apraxia test, we recently provided a clinimetric evaluation in a large cohort of patients with stroke and age-matched healthy subjects.[1] The AST allows for a rapid assessment of apraxia, as the number of items, reduced from 48 to 12, is considerably lower than in the TULIA. Furthermore, the original six-point scoring method was dichotomised to fail and pass so that an 'online' judgement has become possible. The AST can be administered without video recordings which are required for differential

**Table 2** Mean (SD) scores of apraxic (n=19) and non-apraxic (n=12) group

| | AST | | | TULIA | | |
|---|---|---|---|---|---|---|
| | Total | Imitation | Pantomime | Total | Imitation | Pantomime |
| Apraxic | 4.3 (2.9) | 1.5 (1.8) | 2.5 (1.6) | 126.3 (53.5) | 72.8 (26.7)* | 53.5 (31.2)* |
| Non-apraxic | 10.4 (1.0) | 6.3 (0.9)* | 4.2 (0.8)* | 210.6 (9.2) | 106.3 (5.3) | 104.3 (7.2) |

Factorial mixed design ANOVA demonstrated a significant interaction effect between group (apraxic vs non-apraxic) and test (Test for Upper-Limb Apraxia (TULIA) vs Apraxia Screen of TULIA (AST)) (F(1,29)=24.58, p<0.001) for total scores. A significant interaction effect was also found between test (TULIA vs AST) and domain (imitation vs pantomime) (F(1,29)=6.47, p=0.02). *Post hoc paired t tests revealed significant differences (p<0.01) between imitation and pantomime subscores of the AST for the non-apraxic group and of the TULIA for the apraxic group. There was no significant age difference between the apraxic and non-apraxic group (64.1 (16.2) and 62.3 (8.8), p>0.05). Maximum score of AST=12, maximum score of TULIA=240.

**Table 3** Specificity, sensitivity and predictive values of AST and Test for Upper-Limb Apraxia by TULIA

| | | | Test for Upper-Limb Apraxia | | | | |
|---|---|---|---|---|---|---|---|
| | | | | Yes | | | |
| | Apraxia | | No | Mild | Severe | Total | |
| Apraxia Screen of Test for Upper-Limb Apraxia | No | | 11 | 1 | 0 | 12 | Negative predictive value=92% |
| | Yes | Mild | 0 | 12 | 0 | 12 | Positive predictive value=100% |
| | | Severe | 0 | 0 | 7 | 7 | |
| | Total | | 11 | 13 | 7 | 31 | |
| | | | Specificity=100% | Sensitivity=95% | | | |

95% CI: specificity (68 to 100%), sensitivity (73 to 99%), negative predictive value (60 to 99%), positive predictive value (79 to 100%).

scoring schemes but not easily accommodated to bedside testing. The validation with the TULIA revealed the AST to be highly specific and sensitive. The AST proved to be of diagnostic relevance not only in assessing the presence of apraxia reliably within a few minutes, but also in providing a valid classification of the severity of the apraxia, as shown by the almost perfect categorical match between the AST and the TULIA scores. The lower negative predictive value suggests that mild forms of apraxia may be missed with the AST, probably due to the lower sensitivity of the dichotomous scoring method. Likewise, lower sensitivity of the AST may also account for the significant group and test interaction for total scores (table 2). Finally, the high correlation between the AST and the corresponding recoded items from TULIA were indicative of a high test–retest reliability.

Most patients with a left-hemispheric stroke and apraxia are aphasic.[13 14] The strong correlation of the total AST scores with the Token test was not surprising. However, the presence of an aphasia is unlikely to account for the diagnosis of apraxia in our patient cohort, as a strong association was found between the total AST score and the Token test for the imitation part, which is basically independent of language processing. Nonetheless, aphasia with severe comprehension deficits may lead to confounds in praxis testing when it is a matter of performing pantomines on verbal command. Thus, it is conceivable that severely aphasic patients may fall below the cut-off score for apraxia in the AST due to impaired auditory comprehension in the pantomime part. In selected cases, additional testing of the TULIA with separate cut-off scores for imitation may be needed to confirm the diagnosis of apraxia.

A limitation of the study is the possible bias by the examiner's global impression from TULIA on his subsequent scoring of the AST, although the high correlation at the item level would not have been expected based merely on a vague impression. Another concern is the small sample size used for validation; it limits a generalisation of the findings. Finally, the AST has to be validated externally with another standardised test for apraxia in a separate study.

In conclusion, apraxia may go undetected in acute stroke or early rehabilitation units when sensorimotor and language impairments apparently predominate. AST offers a reliable and valid bedside instrument to accurately and quickly diagnose apraxia, which is important for planning targeted and comprehensive rehabilitation programmes early.

**Acknowledgements** We thank F Antonietti and C Held, for their crucial assistance in evaluating the videotapes. We are also grateful to J Milburn and D Weniger, for editing the manuscript, and H Mattle, for his advice.

**Competing interests** None.

**Patient consent** Obtained.

**Ethics approval** Ethics approval was provided by the University of Leuven and University of Bern.

**Provenance and peer review** Not commissioned; externally peer reviewed.

## REFERENCES

1. **Vanbellingen T,** Kersten B, Van Hemelrijk B, *et al*. Comprehensive assessment of gesture production: a new test of upper limb apraxia (TULIA). *Eur J Neurol* 2010;**17**:59–66.
2. **De Renzi E,** Motti F, Nichelli P. Imitating gestures. A quantitative approach to ideomotor apraxia. *Arch Neurol* 1980;**37**:6–10.
3. **Hanna-Pladdy B,** Heilman KM, Foundas AL. Cortical and subcortical contributions to ideomotor apraxia: analysis of task demands and error types. *Brain* 2001;**124**:2513–27.
4. **Ochipa C,** Rothi LJG. Limb apraxia. *Semin Neurol* 2000;**20**:471–8.
5. **Donkervoort M,** Dekker J, Deelman B. The course of apraxia and ADL functioning in left hemisphere stroke patients treated in rehabilitation centres and nursing homes. *Clin Rehabil* 2006;**20**:1085–93.
6. **Foundas AL,** Macauley BL, Raymer AM, *et al*. Ecological implications of limb apraxia: evidence from mealtime behavior. *J Int Neuropsychol Soc* 1995;**1**:62–6.
7. **Hanna-Pladdy B,** Heilman KM, Foundas AL. Ecological implications of ideomotor apraxia: evidence from physical activities of daily living. *Neurology* 2003;**60**:487–90.
8. **Walker CM,** Sunderland A, Sharma J, *et al*. The impact of cognitive impairment on upper body dressing difficulties after stroke: a video analysis of patterns of recovery. *J Neurol Neurosurg Psychiatry* 2004;**75**:43–8.
9. **Goldenberg G,** Hagmann S. Therapy of activities of daily living in patients with apraxia. *Neuropsychol Rehab Int J* 1998;**8**:123–41.
10. **Buxbaum LJ,** Haaland KY, Hallett M, *et al*. Treatment of limb apraxia: moving forward to improved action. *Am J Phys Med Rehabil* 2008;**87**:149–61.
11. **Sunderland A,** Shinner C. Ideomotor apraxia and functional ability. *Cortex* 2007;**43**:359–67.
12. **Poizner H,** Clark MA, Merians AS, *et al*. Joint coordination deficits in limb apraxia. *Brain* 1995; **118**(Pt 1):227–42.
13. **Goldenberg G,** Hartmann K, Schlott I. Defective pantomime of object use in left brain damage: apraxia or asymbolia? *Neuropsychologia* 2003;**41**:1565–73.
14. **Goodglass H,** Kaplan E. Disturbance of gesture and pantomime in aphasia. *Brain* 1963;**86**:703–20.
15. **Almeida QJ,** Black SE, Roy EA. Screening for apraxia: a short assessment for stroke patients. *Brain Cogn* 2002;**48**:253–8.
16. **Novick MR.** The axioms and principal results of classical test theory. *J Math Psychol* 1966;**3**:1–18.
17. **Oldfield RC.** The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 1971;**9**:97–113.
18. **Gauthier L,** Dehaut F, Joanette Y. The Bells Test: a quantitative and qualitative test for visual neglect. *Int J Clin Neuropsychol* 1989;**11**:49–54.
19. **Lang C,** Dehm A, Dehm B, *et al*. *Kurze Aphasieprüfung KAP*. Frankfurt: Swets & Zeitlinger, 1999.

# A new bedside test of gestures in stroke: the apraxia screen of TULIA (AST)

T Vanbellingen, B Kersten, A Van de Winckel, et al.

*J Neurol Neurosurg Psychiatry* 2011 82: 389-392 originally published online October 9, 2010
doi: 10.1136/jnnp.2010.213371