

1-1-2020

## Development and diagnostic validation of the Brisbane Evidence-Based Language Test

Alexia Rohde

Suhail A. Doi

Linda Worrall

Erin Godecke

*Edith Cowan University, e.godecke@ecu.edu.au*

Anna Farrell

*See next page for additional authors*

Follow this and additional works at: <https://ro.ecu.edu.au/ecuworkspost2013>



Part of the [Medicine and Health Sciences Commons](#)

---

[10.1080/09638288.2020.1773547](https://doi.org/10.1080/09638288.2020.1773547)

Rohde, A., Doi, S. A., Worrall, L., Godecke, E., Farrell, A., O'Halloran, R., ... & Wong, A. (2020). Development and diagnostic validation of the Brisbane Evidence-Based Language Test. *Disability and rehabilitation, Advanced online publication*. <https://doi.org/10.1080/09638288.2020.1773547>

This Journal Article is posted at Research Online.

<https://ro.ecu.edu.au/ecuworkspost2013/8392>

---

**Authors**

Alexia Rohde, Suhail A. Doi, Linda Worrall, Erin Godecke, Anna Farrell, Robyn O'Halloran, Molly McCracken, Nadin Lawson, Rebecca Cremer, and Andrew Wong



## Development and diagnostic validation of the Brisbane Evidence-Based Language Test

Alexia Rohde, Suhail A. Doi, Linda Worrall, Erin Godecke, Anna Farrell, Robyn O'Halloran, Molly McCracken, Nadine Lawson, Rebecca Cremer & Andrew Wong

To cite this article: Alexia Rohde, Suhail A. Doi, Linda Worrall, Erin Godecke, Anna Farrell, Robyn O'Halloran, Molly McCracken, Nadine Lawson, Rebecca Cremer & Andrew Wong (2020): Development and diagnostic validation of the Brisbane Evidence-Based Language Test, Disability and Rehabilitation, DOI: [10.1080/09638288.2020.1773547](https://doi.org/10.1080/09638288.2020.1773547)

To link to this article: <https://doi.org/10.1080/09638288.2020.1773547>



© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



Published online: 22 Jun 2020.



[Submit your article to this journal](#)



Article views: 1909



[View related articles](#)



[View Crossmark data](#)

## Development and diagnostic validation of the Brisbane Evidence-Based Language Test

Alexia Rohde<sup>a,b</sup>, Suhail A. Doi<sup>c</sup>, Linda Worrall<sup>b</sup> , Erin Godecke<sup>d</sup>, Anna Farrell<sup>e</sup>, Robyn O'Halloran<sup>f</sup> , Molly McCracken<sup>b</sup>, Nadine Lawson<sup>g</sup>, Rebecca Cremer<sup>e</sup> and Andrew Wong<sup>h</sup>

<sup>a</sup>Speech Pathology Department, Southern Cross University, Bilinga, Australia; <sup>b</sup>Speech Pathology Department, The University of Queensland, Brisbane, Australia; <sup>c</sup>Department of Population Medicine, Qatar University, Doha, Qatar; <sup>d</sup>Department of Speech Pathology, Edith Cowan University, Joondalup, Australia; <sup>e</sup>Department of Speech Pathology, Royal Brisbane and Women's Hospital, Brisbane, Australia; <sup>f</sup>Department of Community and Clinical Allied Health, La Trobe University, Melbourne, Australia; <sup>g</sup>Department of Speech Pathology, Princess Alexandra Hospital, Brisbane, Australia; <sup>h</sup>Department of Neurology, Royal Brisbane and Women's Hospital, Brisbane, Australia

### ABSTRACT

**Purpose:** To describe the development and determine the diagnostic accuracy of the Brisbane Evidence-Based Language Test in detecting aphasia.

**Methods:** Consecutive acute stroke admissions ( $n=100$ ; mean = 66.49y) participated in a single (assessor) blinded cross-sectional study. Index assessment was the ~45 min Brisbane Evidence-Based Language Test. The Brisbane Evidence-Based Language Test is further divided into four 15–25 min Short Tests: two Foundation Tests (severe impairment), Standard (moderate) and High Level Test (mild). Independent reference standard included the Language Screening Test, Aphasia Screening Test, Comprehensive Aphasia Test and/or Measure for Cognitive-Linguistic Abilities, treating team diagnosis and aphasia referral post-ward discharge.

**Results:** Brisbane Evidence-Based Language Test cut-off score of  $\leq 157$  demonstrated 80.8% ( $LR+ = 10.9$ ) sensitivity and 92.6% ( $LR- = 0.21$ ) specificity. All Short Tests reported specificities of  $\geq 92.6\%$ . Foundation Tests I (cut-off  $\leq 61$ ) and II (cut-off  $\leq 51$ ) reported lower sensitivity ( $\geq 57.5\%$ ) given their focus on severe conditions. The Standard (cut-off  $\leq 90$ ) and High Level Test (cut-off  $\leq 78$ ) reported sensitivities of  $\geq 72.6\%$ .

**Conclusion:** The Brisbane Evidence-Based Language Test is a sensitive assessment of aphasia. Diagnostically, the High Level Test recorded the highest psychometric capabilities of the Short Tests, equivalent to the full Brisbane Evidence-Based Language Test. The test is available for download from [brisbanetest.org](http://brisbanetest.org).

### ARTICLE HISTORY

Received 18 January 2020  
Revised 19 May 2020  
Accepted 20 May 2020

### KEYWORDS

Aphasia; diagnosis; test; stroke; sensitivity and specificity; evidence-based

### ► IMPLICATIONS FOR REHABILITATION

- Aphasia is a debilitating condition and accurate identification of language disorders is important in healthcare.
- Language assessment is complex and the accuracy of assessment procedures is dependent upon a variety of factors.
- The Brisbane Evidence-Based Language Test is a new evidence-based language test specifically designed to adapt to varying patient need, clinical contexts and co-occurring conditions.
- In this cross-sectional validation study, the Brisbane Evidence-Based Language Test was found to be a sensitive measure for identifying aphasia in stroke.

## Introduction

Accurate identification of acquired language disorders (aphasia) is important in healthcare. Clinical assessment of language however is complex and multifaceted [1,2] and the accuracy of assessment procedures is dependent upon a number of factors. Performance on language tasks may be easily impacted by multiple other clinical elements which can spuriously influence the accuracy of test scores [2,3]. Co-occurring apraxia of speech or dysarthria may influence verbal language performance and lead to false positives on assessment tasks [3]. Visual deficits such as hemianopia or neglect may interfere with the perception of written language

stimulus or the naming of pictures [3,4] and motor limb impairments such as hemiparesis may impact on writing legibility [5]. Additional factors such as patient fatigue, medical instability and the pressured demands of different clinical contexts also require language testing to be time-efficient, user-friendly and adaptable to the varying needs of the testing environment [2].

A wide range of language tests currently exist. Brief aphasia screening tests such as the Frenchay Aphasia Screening Test (FAST) [6], Language Screening Test (LAST) [7] and ScreeLing [8] aim to provide a fast, efficient gauge of language functioning. These “non-specialist” screeners are intended for use by multiple professions and focus on efficiency and brevity in aphasia

CONTACT Alexia Rohde  [alexia.rohde@scu.edu.au](mailto:alexia.rohde@scu.edu.au)  Department of Speech Pathology, Southern Cross University, Bilinga, QLD 4225, Australia

© 2020 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives License (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited, and is not altered, transformed, or built upon in any way.

screening. Such tests however gauge functioning through a narrow range of language tasks, do not assess language ability across the severity spectrum from severe to mild, and at times omit assessment of core language domains such as reading and/or writing [7]. As a result, any deficits within any these areas may be missed.

Comprehensive language tests such as the Western Aphasia Battery-Revised (WAB-R) [3], Comprehensive Aphasia Test (CAT) [9] and Boston Diagnostic Aphasia Examination (BDAE) [10] provide a thorough evaluation of language across the domains. Such measures however take up to 1-2 h to administer and have been reported to be too cumbersome for use in certain settings (e.g., acute hospital sessions which typically last ~30 min) [2,11]. Other, more time-efficient speech pathology measures such as the Acute Aphasia Screening Protocol [12] and WAB-R Bedside [3] are widely used in practice. However, surprisingly, systematic reviews of the literature have failed to find any published studies reporting these tests' validation in stroke patients with and without aphasia [13,14].

The complexity of clinical language evaluations highlights the need for an evidence-based, adaptable, time-efficient yet comprehensive language test. The Brisbane EBLT is a new language test created to address this need. The test has been developed from Evidence-Based Principles (EBP) [15] and is available in five test versions specifically developed to accurately assess language while adjusting to patient need, the clinical context and co-occurring conditions. The aim of this study is to report on the development and diagnostic validation of the Brisbane EBLT in identifying aphasia in stroke populations.

## Materials and methods

### Study design and participants

Diagnostic accuracy was examined in a single (assessor) blinded cross-sectional study of 100 consecutive acute stroke admissions at two large tertiary hospitals in Brisbane, Australia. A priori sample size calculations revealed 98 patients were required to yield a 10% width of a 2-sided 95% CI with a conservative estimate of 50% disease prevalence [16] (estimates of aphasia in left-hemisphere brain damage reported to be 55.1%) [17]. The project was granted university and health service committee ethics approval and aims to comply with recognised Standards for Reporting Diagnostic Accuracy (STARD) [18]. The STARD statement is an EQUATOR network guideline (Enhancing the QUALity and Transparency Of health Research) of widely accepted criteria for the rigorous reporting of sample selection, study design and statistical analysis in diagnostic accuracy research [18].

### Target condition

The target disorder was defined as impaired language functioning from severe to mild deficits within any language domain (verbal expression, comprehension, reading, writing, and gesture) resulting from ischaemic or haemorrhagic stroke.

### Inclusion/exclusion criteria

All admitted stroke patients were screened for eligibility within 2 days of hospital admission. Eligible patients were those admitted for stroke management and deemed sufficiently medically stable to undergo language assessment if the following were met: sustained level of consciousness for >10 min (cognitive functioning was pragmatically assessed based on a patient's ability to

participate in, engage with and complete the required language tasks); absence of any precluding acute medical condition as per treating medical team; aged >14 years; native-level English language ability, and with confirmed stroke site of lesion (as determined by acute medical CT/MRI report) within the left frontal, parietal, temporal, occipital, limbic or insular lobes, basal ganglia (caudate nucleus, putamen, globus pallidus, substantia nigra, nucleus accumbens, and subthalamic nucleus), internal capsule, and thalamus (including thalamic nuclei). To optimise test external validity, the presence of common post-stroke non-language but communication-related conditions (affecting vision, hearing, speaking, or writing) such as hemianopia, hemiparesis, dysarthria or apraxia of speech was not used as an exclusionary criterion. For these patients, the presence of these co-occurring conditions was noted, and language test items severely affected by these conditions (such that they were uninterpretable) were recorded as missing data. Patients with subarachnoid haemorrhage or lesions isolated to the right cerebral hemisphere, right subcortical regions or midbrain, or below (as identified by acute medical CT/MRI report) were excluded [19].

## Assessments

### Brisbane EBLT test development

The Brisbane EBLT is a new language test conceived and created by the first author, Alexia Rohde. The first phase of test development involved the construction of a large item bank of stimulus items including both language tasks (e.g., repetition, reading aloud, picture-word matching) and questions (e.g., picture naming "what is this called?"). The aim of the item bank was to develop as many new tasks and questions as possible for the new assessment. This item bank was guided by the structure and content of existing informal language measures ( $n=44$ ) which were collected from clinicians *via* an email request sent through a professional speech pathology network. Based on these informal measures, similar new tasks and questions were created which formed the basis of the new test. These items were created to assess language across all the domains of verbal expression, auditory comprehension, gesture, reading and writing. Linguistic factors, such as word length, stimulus imageability and word frequency were all incorporated into the test development process. Tasks and questions were designed to assess language across the severity spectrum, from simpler (e.g., picture matching tasks) to complex high-level questions (e.g., synonyms "what is one word that means the same as...?"). Item bank tasks and questions were categorised into a specific area (language domain and difficulty level) to ensure each language area and severity of language functioning was assessed.

At this phase in the test development, the item bank included over 100 subtests and was far too lengthy for practical use. The next phase involved the shortening of the item bank to a workable test length (~50 subtests). Tasks and questions which were ambiguous, too long, not practical, or difficult to score were eliminated to leave only the most effective, efficient and easy-to-score items within the final measure. During this phase each language task and question underwent a rigorous process of piloting, feedback, revision and review based on the four principles of EBP [15]:

**Clinician experience.** A total of 108 clinicians provided feedback on test items during the development process. Speech pathologists provided feedback in a variety of settings including as part of their attendance at a national conference, as part of focus-groups and during one-to-one test administration sessions.

During this time the test was a dynamic document that underwent frequent revision and refinement. Speech pathologists reviewed the new assessment and then provided feedback. Changes to the test were then made and the test then underwent further speech pathology review. The test became progressively shorter as duplicate/similar, ambiguous and difficult-to-score tasks and questions were excluded. This pattern was repeated until nil new feedback on the document was obtained.

**Clinical context.** Piloting of the test in the acute hospital environment ( $n=10$  patients) led to the exclusion of test items too lengthy or cumbersome for acute use. Subtests became short in length (maximum of 5–6 tasks/questions) to ensure the test moved quickly and swiftly across assessment areas. In the final phase of test development, two speech pathology hospital departments piloted the test in their clinical caseloads ( $n=17$  clinicians) and gave further feedback on test items.

**Patient perspectives.** Directly following administration, item bank feedback from stroke patients and/or family members ( $n=74$ ) was sought. Potential cultural bias was minimised with feedback obtained from participants originating from the following English speaking countries: USA, England, Australia, New Zealand, Scotland, Canada, and South Africa. Any test items with reported cultural bias (items unfamiliar to participants from any cultural background) were excluded. The majority of patients ( $n=57$ ) had positive/nil concerns with the new test (“I enjoyed it,” “Was what I expected it,” “No worries,” etc.). Feedback relating to specific tasks or questions was provided by three study participants/their family members which was incorporated into the test development process. For example, the parent of one study participant reported that “Some task instructions were quite complex however the tasks were low. Like the ‘semantic’ ones.” For simpler items, it was subsequently ensured that an example was always first completed by the administering clinician to demonstrate the type of required response. Verbal understanding of the instructions is therefore not needed.

Five of the study participants reported the test was “too long” ( $n=5$ ), “too simplistic” ( $n=5$ ) or “too difficult” ( $n=4$ ). To minimise patient fatigue, avoid patient exposure to test items considered too difficult or easy and replicate the typical length of existing informal measures, the Brisbane EBLT was then subsequently also made available in four shorter versions (Short Tests). Clinician experience, clinical context and patient feedback were all instrumental in guiding the test development process. Only tasks and questions meeting these rigorous EBP standards were included in the final measure, the Brisbane EBLT.

**Clinically relevant research.** The final EBP principle is the focus of this report which describes the diagnostic analysis of this new test.

### **Brisbane EBLT**

The above test development process resulted in the final measure, the Brisbane EBLT which is the full version of the assessment, evaluating language in 49 subtests (~45 min) across the severity spectrum (severe to mild) in the following language domains (Table 1): verbal expression including repetition, automatic speech, spontaneous speech (picture description), naming, auditory comprehension, actions/gesture, reading, and writing. Certain subtests require the use of two of each of the following everyday objects: cup, spoon, pen and knife. An additional “Perceptual” subtest examines abilities not requiring a verbal or written

response (e.g., object to picture matching) to allow assessment of functioning when both verbal and written responses may be difficult (for example in situations when both severe apraxia of speech may result in limited verbal expression and dominant-hand hemiparesis may limit written responses). Large print/pictures and the use of everyday objects as stimulus items assist to accommodate visual co-occurring conditions (e.g., hemianopia). Short, concise writing tasks (for example “Write ONE sentence which contains ALL of these words in THIS order” requires only a single sentence written response) minimise upper limb motor fatigue associated with hemiparesis or non-dominant hand use. The 49 Brisbane EBLT subtests and description of test items are listed in Table 1.

### **Test administration and scoring**

The Brisbane EBLT does not have a user’s manual. All information required to use and administer the test is contained on the test forms themselves and the Administration and Scoring Guidelines form (brisbanetest.org). Question-specific administration and scoring information is contained on the Brisbane EBLT test forms whereas overall scoring guidelines are provided on the Administration and Scoring Guidelines form. The same Administration and Scoring Guidelines apply to all test versions.

Test administration is standardised. Each question is administered by following the task instructions given above each question. Prompting (e.g., the provision of an initial phoneme or additional gestural cue) is not allowed. All tasks must be administered according to the set instructions. One repetition of task instructions is allowed without penalty for non-language related difficulties in interpreting task instructions (e.g., hearing impairment). This repetition has no influence on scoring (e.g., if the patient answers correctly on the second (louder) repetition, the response is still scored as correct).

Patient responses should be scored according to the first purposeful response within the target modality (e.g., the first purposeful attempt at a verbal response is scored, any prior unrelated vocalisations are not scored). The most common responses are provided on the test form. If an alternative yet correct response is given (other than those listed), as judged by the administering clinician, this is still scored as correct. Scores are not deducted for responses impacted by motor related output deficits (e.g., phonetic distortions due to dysarthria or reduced writing legibility due to hemiparesis) providing the target is still interpretable. If the response is not interpretable (e.g., illegible writing or unintelligible speech), these items are to be left blank with a brief note reporting the reason for the lack of score. Non-language related self-correction (e.g., attempts to neaten messy writing due to hemiparesis) are not penalised (still scored as correct), however if any language-related self-corrections are evident (e.g., self-correcting phonemic or semantic paraphasias), these are scored incorrect. All Brisbane EBLT test forms and the Administration and Scoring Guidelines are available from brisbanetest.org.

### **Language reference standard**

The reference standard diagnosis was based on five clinical factors; the results from a battery of published tests interpreted in the context of four additional clinical factors [21]:

1. A composite language test score: All patients were first administered the Language Screening Test (LAST) [7] a brief (2 min) “non-specialist” screening test which provided an initial gauge of language functioning and directed which comprehensive language battery would be most appropriate.

Table 1. Brisbane EBLT subtests.

Subtest	Task description	Score	Total
<b>Perceptual</b>			
1. Copying Gestures	"Copy what I'm doing" (e.g., point to nose).	/2	
2. Object to Object Matching <sup>a</sup>	"Which object matches with this?" (e.g., matching cups)	/2	
3. Demonstrating Object Use <sup>a</sup>	"Show me how you would use these objects" (e.g., pen, writing).	/2	
4. Demonstrating Gestures from Pictures	"Show me the action that goes with each picture" (e.g., toothbrush).	/2	
5. Object to Picture Matching <sup>a</sup>	"Match the object with the right picture" (e.g., spoon).	/2	
6. Picture to Picture Matching (semantic links)	"Which of the bottom two pictures goes best with the middle one?" (e.g., matching apple with banana)	/5	/15
<b>Auditory comprehension</b>			
7. Yes/No Questions	"Answer these questions" (e.g., Is it night time?)	/12	
8. Following Verbal Commands	"I'd like you to do what I say" (1–4 stage commands).	/8	
9. Identifying Pictures by Description	"Point to the ... ." (e.g., boat).	/6	
10. Identifying Objects by Function	"I would like you to point to some different body parts" (e.g., point to where you would wear a hat).	/2	
11. Odd-one-out	"Tell me, which one of these words is the odd one out or doesn't belong." (4 words) "Why?"	/2	
12. Complex Question	"Answer these questions" (e.g., repeat these words in alphabetical order (4 words)).	/6	
13. Self-Reported Difficulty?	"Do you think these last few questions are harder for your now since you have had your stroke?"	Yes/No	
14. Synonyms	What is one word that means the same as ... ?" (e.g., surprised/shocked).	/2	/38
<b>Verbal expression<sup>c</sup></b>			
15. Automatic Speech	"Count from 1 to 10."	/1	
16. Sentence Completion	"Finish what I say" (e.g., Boys and ... ).	/2	
17. Personal/Orientation Questions	"Tell me, what is?" (e.g., your name?)	/6	
18. Repetition	"Repeat what I say" (3 words, 1 sentence).	/4	
19. Object Naming <sup>a</sup>	"What is this called?" (e.g., knife).	/2	
20. Naming Actions <sup>a</sup>	"What do you call this action?" (e.g., drinking).	/2	
21. Picture Naming	"Look at these pictures. What is this?" (e.g., car).	/4	
22. Naming Objects from the Room <sup>b</sup>	"What's this called?" (e.g., pillow).	/4	
23. Naming Gestures	"Tell me what action I'm doing" (e.g., salute).	/2	
24. Verbal Fluency – Animals	"Name as many different animals as you can in 1 minute."	/45 <sup>d</sup>	
24. Verbal Fluency – Words Starting with F	"Name as many words starting with the letter "F" as you can in 1 minute."	/35 <sup>d</sup>	
25. Picture Description	"Look at this picture, it tells a story. Tell me what you see." (stimulus is the Cookie Theft Picture [10] and scoring based on no. of content units produced) (scoring system adapted from Yorkston & Beukelman [20]).	/16	
26. Self-Reported Difficulty?	"Did you find that task difficult?"	Yes/No	
27. Word Definitions	"Tell me what these words mean. Explain the word as fully as you can" (e.g., school).	/4	
28. Similarities and Differences	"Tell me what is similar and what is different about these things."	/2	
29. Proverbs	"Tell me, what does this saying mean?"	/3	/132
<b>Reading</b>			
30. Object to Word Matching <sup>a</sup>	"Match the object with the right word." (e.g., spoon).	/2	
31. Single Word Reading	"Read these words and then point to the body part" (e.g., ear).	/2	
32. Written Word to Picture Matching	"Read each of these words aloud and then point to the right picture" (e.g., diamond).	/6	
33. Following Written Commands	"Read this and then do what it says" (1–2 stage commands).	/4	
34. Sums	"Solve these" (calculations).	/2	
35. Reading Aloud	"Read this aloud" (sentence).	/1	
36. Medicine Label	"Here is a medicine label" (e.g., How many tablets per day?)	/3	
37. High Level Sentence Comprehension	"Read these instructions and answer the questions" (complex questions).	/2	
38. Self-Reported Difficulty?	"Did you feel you had any difficulty reading and understanding the words as you read?" (reading comprehension).	Yes/No	
39. Delayed Recall <sup>c</sup>	"Now, tell the exact story back to me." (written paragraph recall).	/21	
40. Inference	"What do you think really happened?"	/1	/44
<b>Writing</b>			
41. Drawing Completion	"Look at this picture. Draw in the part that is missing."	/1	
42. Simple Copying	"Look at these. Copy them below."	/2	
43. Functional Writing – Name	"Write your full name here."	/2	
44. Functional Writing – Address	"Please fill in the form."	/3	
45. Writing to Dictation	"Write down what I say."	/6	
46. Written Naming – Object	"I'm going to point to something. Write down what it's called" (index finger.)	/1	

(continued)

Table 1. Continued.

Subtest	Task description	Score	Total
47. Written Naming – Gesture	“Watch what I am doing, and then write it down” (clapping).	/1	
48. Sentence Construction	“Write one sentence which contains all of these words”	/13	
49. Self-Reported Difficulty?	“Do you think you found it harder than you would have before your stroke?”	Yes/No	/29
<b>Total Brisbane EBLT Score:</b>		<b>/258</b>	
Adapted score: excluding hospital ward items <sup>b</sup>		/254	
Adapted score: excluding verbal expression subtests <sup>c</sup>		/105	

<sup>a</sup>Requires everyday objects. Test items require two of each of the following everyday objects (cup, spoon, pen and knife).

<sup>b</sup>Adapted score: excluding hospital ward items. Enables the test to be administered in a clinic or non-hospital ward location. Total score is calculated by omitting the subtest requiring the naming of hospital ward objects (bed, pillow etc).

<sup>c</sup>Adapted score: excluding verbal expression subtests. Enables the test to be administered with patients with co-occurring expressive deficits affecting their verbal responses (e.g., apraxia of speech). Total score is calculated by omitting the Verbal Expression section (subtests 15–29) & Delayed Recall (subtest 39) and summing performance from the remaining four modalities (auditory comprehension, perception, reading and writing).

<sup>d</sup>Score of/45 and/35 are maximum possible scores only (not cut-off scores). High maximum scores of/45 and/35 were selected to minimise the likelihood any patient will achieve them and consequently all patient responses are captured in the test total.

Patients were then administered one of three longer comprehensive batteries, the Aphasia Screening Test (AST) [22] (LAST scores  $\leq 5/15$ ), the Comprehensive Aphasia Test (CAT) [9] (LAST scores of 5–15), or the Measure for Cognitive-Linguistic Abilities (MCLA) [23] (LAST scores  $\geq 10$ ).

2. The treating clinician’s documented diagnosis,
3. Multidisciplinary report
4. Medical team report, and
5. Presence of referral for language services post-discharge

Reference diagnosis was based on a majority of clinical determinants of  $\geq 3$  out of the five areas.

### Procedure

Consecutive stroke admissions from 21 January to 15 December 2015 were reviewed for eligibility. Informed written consent was obtained for all participants or an officially authorised next of kin. Recruited participants completed the full 49 subtest Brisbane EBLT and reference measure randomised in order. Brisbane EBLT sections were also randomised for each participant. The assessment procedure was single (assessor) blinded. Assessments were administered as closely together as possible at the patient’s bedside or ward clinic room by two speech pathologists blinded to the other’s test results. Study participants were not directly informed of their aphasia diagnosis status as part of their participation in this study, however their diagnosis status likely would have been disclosed by members of their treating multidisciplinary care team as part of their routine acute stroke care.

### Statistical analysis

Diagnostic analysis was completed using MedCalc v 13.3.3.0 and/or Stata/IC 13.0. Diagnostic accuracy was calculated by comparing Brisbane EBLT test scores against the binary (yes/no) reference result to determine test sensitivity, specificity, positive (+LR), and negative likelihood ratios (–LR). Test discrimination against the reference standard was determined using the area under (AUC) the receiver operating characteristic (ROC) curve. Cut-off scores for discriminating between patients with and without aphasia were determined by evaluating the sensitivity and specificity estimates at each cut-off threshold score.

### Brisbane EBLT Short Tests

The Brisbane EBLT was administered to the 100 patients for the purposes of data collection to allow for a full dataset of all 49 subtests of the Brisbane EBLT to be captured. However, at

~45 min in length this full test version of the Brisbane EBLT is inappropriate for use in many contexts (e.g., routine acute hospital bedside use) [2,11]. To replicate the <30 min length of existing informal acute language measures and allow adaption to varying contexts, the Brisbane EBLT dataset was divided to create four Short Tests (15–25 min each).

Each Brisbane EBLT Short Test assesses language across the domains of verbal expression, auditory comprehension, reading and writing yet focuses on a particular severity level (mild, moderate or severe). Two Short Tests focus on severe language deficits: 1) “Foundation Test I” (with objects) (cup, spoon, pen and knife) and 2) “Foundation Test II” (requiring nil objects). The “Standard Test” examines moderate deficits and the “High Level Test” identifies mild/nil impairment. Assessment items from the full Brisbane EBLT were broadly allocated to a Short Test according to the percentage of participants who answered each question correctly (>75% for the Foundation Tests, 75–35% for the Standard Test and >45% for the High Level Test). Based on clinician feedback some questions were included within two test versions (e.g., Picture Description is within the Standard and High Level Tests). Short Test subtests are listed in Table 2.

To ensure clinical applicability of these new shorter measures, all Short Test underwent acute hospital piloting and speech pathology feedback to ensure the tests were effective in real-world clinical contexts. Diagnostic analysis was separately completed for all Short Tests allowing each to act as standalone assessments enabling clinicians to select any test version at their discretion based on their initial clinical impression, the context, the presence of any co-occurring deficits and level of language functioning they wish to assess (thereby preventing patients from being exposed to tasks markedly too simple/difficult). The same scoring method applies to all test versions where each question is attributed a score adding to a total score used to indicate the presence of aphasia. Full scoring instructions are available at [brisbanetest.org](http://brisbanetest.org).

### Adapted test scores

Additional subdivision of the Brisbane EBLT datasets was conducted to create adapted test scores. Adapted scores enable the tests to be administered and total scores calculated despite the elimination of certain subtests which may be required depending upon different testing environments. The full Brisbane EBLT and Short Tests (excluding High Level) have adapted test versions to adjust to the context (e.g., limited time), patient ability (e.g., fatigue) or when severe co-occurring conditions are present (e.g., hemiparesis, hemianopia, dysarthria or apraxia of speech). The full Brisbane EBLT has two adapted scores. The first excludes the



Table 2. Brisbane EBLT Short Tests.

Foundation Test I (with objects) <sup>a</sup>		Foundation Test II		Standard Test		High Level Test	
PERCEPTUAL		PERCEPTUAL		AUDITORY COMPREHENSIONS		AUDITORY COMPREHENSION	
1. Copying Gestures	/2	1. Copying Gestures	/2	1. Yes/No Questions	/12	1. Following Verbal Commands	/6
2. Object to Object Matching	/2	2. Demonstrating Object Use from Pictures	/2	2. Following Verbal Commands	/6	2. Odd-one-out	/2
3. Demonstrating Object Use	/2	3. Picture to Picture Matching	/5			3. Complex Question	/6
4. Demonstrating Object Use from Pictures	/2			<b>VERBAL EXPRESSIONS</b>		4. Self-Reported Difficulty?	Y/N
5. Object to Picture Matching	/2	<b>AUDITORY COMPREHENSION</b>		3. Automatic Speech	/1	5. Synonyms	/2
6. Picture to Picture Matching	/5	4. Yes/No Questions	/12	4. Sentence Completion	/2		
		5. Following Verbal Commands	/4	5. Personal/Orientation Questions	/6	<b>VERBAL EXPRESSION</b>	
<b>AUDITORY COMPREHENSION</b>		6. Identifying Pictures by Description	/6	6. Repetition	/4	6. Personal/Orientation Questions	/6
7. Yes/No Questions	/12	7. Identifying Objects by Function	/2	7. Picture Naming	/4	7. Verbal Fluency – Animals	/45
8. Following Verbal Commands	/4			8. Naming Objects from the Room <sup>b</sup>	/4	7. Verbal Fluency – F words	/35
9. Identifying pictures by Description	/6	<b>VERBAL EXPRESSION</b>		9. Naming Gestures	/2	8. Picture Description	/16
10. Identifying Objects by Function	/2	8. Automatic Speech	/1	10. Picture Description	/16	9. Self-Reported Difficulty?	Y/N
		9. Sentence Completion	/2	11. Self-Reported Difficulty?	Y/N	10. Word Definitions	/4
<b>VERBAL EXPRESSION</b>		10. Personal/Orientation Questions	/4	12. Verbal Fluency – Animals	/45	11. Similarities and Differences	/2
11. Automatic Speech	/1	11. Repetition	/3	12. Verbal Fluency – F Words	/35	12. Proverbs	/3
12. Sentence Completion	/2	12. Picture Naming	/2				
13. Personal/Orientation Questions	/4	13. Naming Objects from the Room <sup>b</sup>	/3	<b>PERCEPTUAL</b>		<b>READING</b>	
14. Repetition	/3			13. Picture to Picture Matching	/5	13. Sums	/2
15. Object Naming	/2	<b>READING</b>				14. Reading Aloud	/1
16. Naming Actions	/2	14. Single Word Reading	/2			15. Medicine Label	/3
17. Picture Naming	/2			<b>READING<sup>c</sup></b>		16. High Level Sentence Comp	/2
18. Naming Objects from the Room <sup>b</sup>	/3	<b>WRITING</b>		14. Written Word to Picture Matching	/6	17. Self-Reported Difficulty?	Y/N
		15. Drawing Completion	/1	15. Following Written Commands	/4	18. Delayed Recall	/21
<b>READING</b>		16. Simple Copying	/2	16. Reading Aloud	/1	19. Inference	/1
19. Object to Word Matching	/2	17. Functional Writing	/2	17. Functional Reading – Medicine Label	/3		
20. Single Word Reading	/2					<b>WRITING</b>	
		<b>WRITING</b>				20. Writing to Dictation	/4
		21. Drawing Completion	/1			21. Written Naming – Object	/1
		22. Simple Copying	/2			22. Written Naming – Gesture	/13
		23. Functional Writing – Name	/2			23. Sentence Construction	Y/N
						24. Self-Reported Difficulty?	
<b>Foundation Test I Total Score<sup>a</sup></b>	<b>/67</b>	<b>Foundation Test II Total Score</b>	<b>/55</b>	<b>Standard Test Total Score</b>	<b>/169</b>	<b>High Level Test Total Score</b>	<b>/176</b>
Adapted score: excl hosp ward items <sup>b</sup>	/64	Adapted score: excl hosp ward items <sup>b</sup>	/52	Adapted score: excl hosp ward items <sup>b</sup>	/165		
				Adapted score: excl reading/writing <sup>c</sup>	/137		

<sup>a</sup>Requires everyday objects. Test items require two of each of the following everyday objects (cup, spoon, pen and knife).

<sup>b</sup>Adapted score: excluding hospital ward items. Enables the test to be administered in a clinic or non-hospital ward location. Total score is calculated by omitting the subset "Naming objects from the room" which requires ward objects to be present (bed, pillow etc).

<sup>c</sup>Adapted score: excluding reading/writing. Enables the test to be administered with patients with significant fatigue or co-occurring deficits affecting reading (e.g., hemianopia) and/or written responses (e.g., hemiparesis). Total score is calculated by omitting reading and writing tasks. Note, any dysgraphia and dyslexia will not be identified.

subtest requiring the naming of hospital ward objects (e.g., bed, pillow) and enables the test to be administered in a clinic room (the Foundation Tests and Standard Test also have this adapted version). The second adapted version allows the calculation of a total score while excluding subtests which rely heavily upon verbal expression as the mode of patient response. Excluding these subtests minimises the confounding of the overall total test scores by the presence of severe co-occurring dysarthria or apraxia of speech. The Standard Test has an additional adapted score which excludes reading and writing tasks and enables the test to be administered when there is limited time, patient fatigue or when there are severe co-occurring deficits affecting reading (e.g., hemianopia) and/or written responses (e.g., hemiparesis). Note, any dysgraphia and dyslexia will not be identified using this adapted score. Diagnostic analysis was completed for all adapted test scores. Brisbane EBLT adapted scores are reported in Table 1 and Short Tests adapted scores in Table 2.

#### *Indeterminate Brisbane EBLT and reference standard results*

The Brisbane EBLT was developed with the aim to minimise the influence of non-language related deficits (e.g., hemiparesis, hemianopia, apraxia of speech) on language test scores, however when co-occurring deficits were severe (e.g., hemiparesis which prevented the completion of any writing tasks) affected test scores were recorded as missing and scored zero to ensure resultant scores were as pragmatic and reflective of clinical practice as possible. Reference standard tests were scored according to manual instructions. If reference standard data was missing or impaired due to non-language deficits, subtest scores were used instead. For all participants, multiple factors informed reference standard diagnosis.

## Results

### *Participants*

Of 154 eligible admissions, 120 (77.9%) stroke patients consented to participate. Of these, 100 (64.9%) completed both assessments (Figure 1). Cognition was not screened or assessed formally or informally but rather the participants' successful engagement in and completion of the assessment process made them eligible to participate and have their data used in the protocol. The flow-chart of study eligibility and recruitment is provided in Figure 1. Participant characteristics (patients who completed both assessments) are described in Table 3.

To optimise test external validity, common post-stroke co-occurring conditions were not used as exclusion criteria. Patients with non-language but communication-related co-occurring conditions (affecting vision, hearing, speaking, or writing) were included within the study sample. The Brisbane EBLT was developed to minimise the impact of these conditions (e.g., large stimulus/text size) with the specific aim of enabling the majority of patients to complete assessment items. This test design was found to be able to accommodate the majority of co-occurring conditions, enabling patients who presented with mild or moderate co-occurring deficits (e.g., mild-moderate hemiparesis) to still perform set tasks and questions. However, a limited number of patients who presented with severe co-occurring conditions (e.g., severe hemiparesis) were still unable to complete certain Brisbane EBLT assessment items. The percentage of included patients presenting with co-occurring conditions and the number of patients unable to complete certain tasks due to the severity of their impairment (reported in brackets) were: hearing impairment 43%  $n = 43$  (0 patients) (i.e., 43 (43%) patients reported some form of

hearing impairment however all of these deficits were able to be overcome (e.g., repetition of instructions/increasing volume) in all cases such that 0 patients were unable to complete items due to the severity of their hearing deficits); hemianopia 27.06%  $n = 26$  (4 patients); hemiparesis 60%  $n = 60$  (8 patients); upper limb apraxia 5%;  $n = 5$  (1 patient); apraxia of speech 25%  $n = 24$  (13 patients); and dysarthria 34%  $n = 34$  (4 patients).

Statistically, the inclusion of patients with severe co-occurring conditions had little impact on the overall psychometric dataset. When patients had severe deficits in one area (e.g., severe hemiparesis resulting in inability to write), these patients were often able to complete items in other areas (e.g., verbal tasks such as following commands). As a result, the overall impact of these affected test items was found to constitute <5% of the recorded data and consequently the impact of missing data was considered statistically negligible [24]. In clinical practice however, when patient performance needs to be summed to calculate an overall test score, the creation of Brisbane EBLT Adapted Test Scores assists with further adjusting the test to accommodate the presence of these severe co-occurring conditions.

### *Distribution of severity of language impairment*

Seventy-three (73%) participants were diagnosed with aphasia according to the reference standard. Of these, seven were assessed using the AST [22] (severe aphasia), 36 with the CAT [9] (moderate aphasia), and 57 with the MCLA [23] (mild or nil aphasia). The Brisbane EBLT demonstrated no floor or ceiling effects with scores ranging from 7 to 215 (out of a possible 0 to 258) (Figure 2).

### *Time interval between index test and reference standard*

Patients were admitted to hospital on average 1 day 14 h post-stroke. The average time between admission and completion of both assessments was 7 days 3.75 h (ranging from 1 day 21 h to 28 days 6 h). The Brisbane EBLT took 48.09 min and the reference standard took 76.05 min on average to complete. The average time between assessments was 23 h 39 min. Patients received only dysphagia related speech pathology intervention between the two tests.

### *Adverse events from performing the index test or reference standard*

Patients reported fatigue when completing the reference standard. In the absence of an existing recognised acute diagnostic language test, the reference battery was lengthy (average 76.05 min) and consequently not optimal for acute use [2,11].

### *Estimates of diagnostic accuracy*

#### *Brisbane EBLT*

An overall Brisbane EBLT cut-off score of  $\leq 157$  had a sensitivity of 80.8% (95% CI, 69.9–89.1) and specificity of 92.6% (95% CI, 75.7–99.1) with LR+ 10.9 and LR– 0.21. ROC analysis indicated an AUC of 0.908 (SE 0.030) (95% CI, 0.85–0.97). Threshold scores indicating presence of language impairment were determined based on the cut-off that yielded the highest sensitivity for a specificity  $\geq 90\%$ . Score distributions indicated that aphasia diagnosis represented a spectrum ranging from mild to severe while scores of language-intact stroke patients clustered in a narrower range thus creating a shoulder effect for specificity. Cut-off score distribution

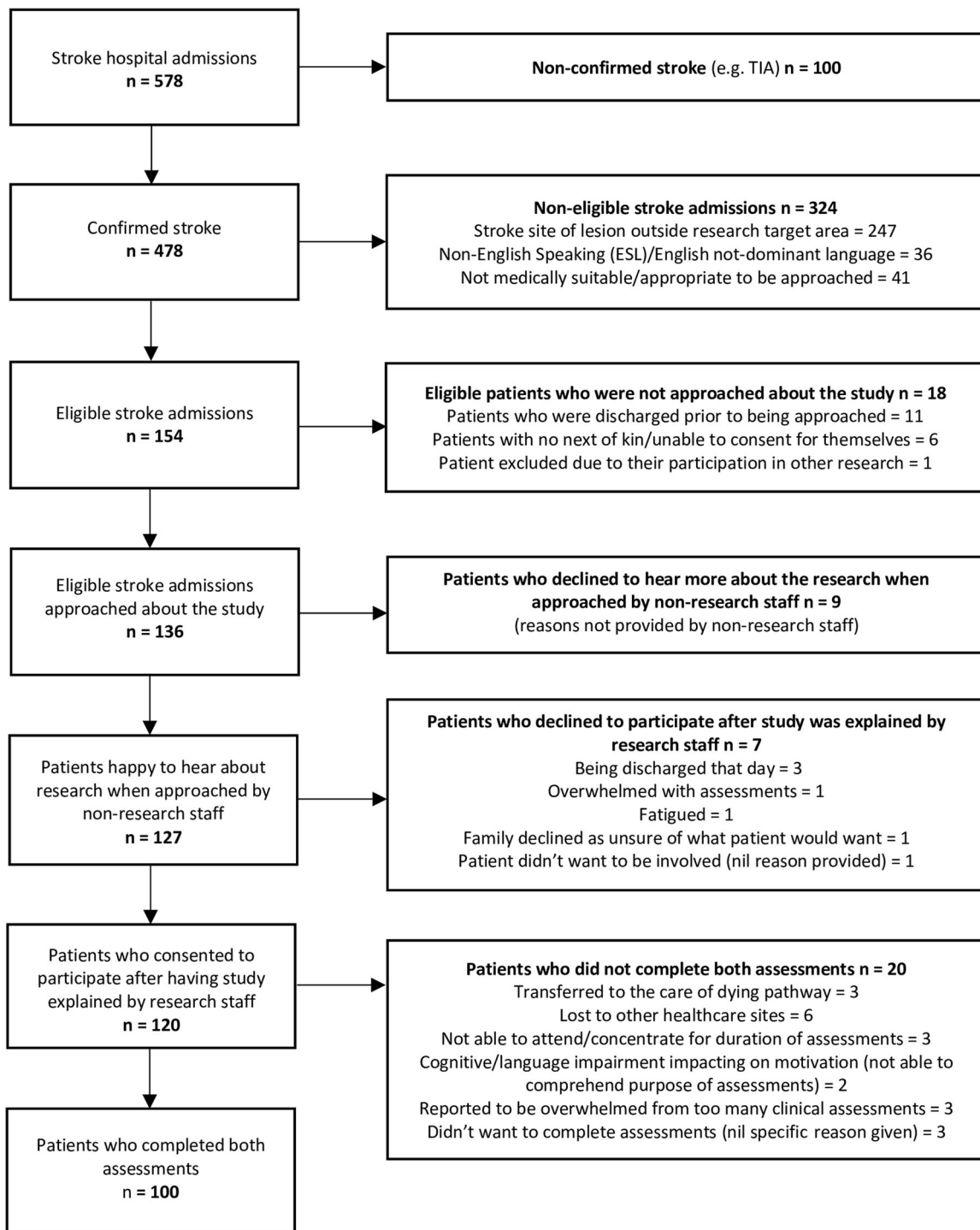


Figure 1. Flowchart of patient eligibility and recruitment.

analysis indicated that despite the lower sensitivity estimates ( $\geq 80.8\%$ ) obtained from this cut-off, patients diagnosed with aphasia who achieved Brisbane EBLT scores of  $\leq 157$  documented a similar score range to those without language impairment (Figure 3). In circumstances when a single (binary) cut-off score may be needed (e.g., for research), these lower cut-off thresholds should be used. However, to assist with clinical exclusion of the condition, threshold scores above which indicate likely absence of impairment were also determined based on higher cut-offs (e.g.,  $\leq 177$  ( $\geq 178$ )) that yielded sensitivities  $\geq 90\%$ . Scores between

the two cut-offs indicated possible risk of impairment. Age, education level, and gender were not found to significantly affect test scores. Cross tabulation of Brisbane EBLT scores by reference standard and diagnostic accuracy estimates (including adapted scores) are listed in Table 4.

#### Short Tests

Short Test versions all reported specificities of  $\geq 92.6\%$  with the Standard and High Level Tests reporting sensitivities of  $\geq 72.6\%$  (excluding adapted scores). Foundation Tests reported lower

sensitivity estimates of  $\geq 57.5\%$  given their focus on more severe conditions. ROC analysis indicated an AUC of  $\geq 0.808$  (SE 0.041) (95% CI, 0.73–0.89) for all Short Tests (Table 4). Diagnostically, the High Level Test reported the highest estimates of the Short Tests, equal to the full Brisbane EBLT with a sensitivity of 80.8% (95% CI 69.9–89.1) and specificity of 92.6% (95% CI 75.7–99.1).

### Discussion

This multicentre cross-sectional study describes the development and diagnostic validation of the Brisbane EBLT. The use of EBP [15] in guiding test development aims to ensure the test meets high psychometric standards and is feasible, user-friendly and adaptable to varying clinical contexts, patient abilities and settings. The Brisbane EBLT demonstrates good sensitivity of 80.8% (95% CI, 69.9–89.1) and a high specificity of 92.6% (95% CI, 75.7–99.1) in identifying aphasia.

Clinically, the 15–25 min Short Tests are most appropriate for use in acute hospital use. Diagnostically, all five test versions (the full Brisbane EBLT and four Short Tests) demonstrate specificity of  $\geq 90\%$  (excluding adapted scores). Of these, the High Level Test

reports the highest diagnostic estimates, equivalent to that of the full Brisbane EBLT.

### Comparison with other research, implications for practice, intended use, and clinical role of the test

A wide range of language measures currently exist to aid in the clinical assessment of aphasia [2]. A number of brief aphasia tests such as the LAST [7] and Quick Aphasia Battery (QAB) [25] report sensitivity and specificity data in stroke populations. While these tests aim to quickly and efficiently assess language functioning in the acute post-stroke environment, these brief tests are either intended for “non-specialist” screening and/or lack assessment across all core language domains of auditory comprehension, verbal expression, reading comprehension and writing [7,25]. Other measures such as the WAB-R [3], CAT [9] and MCLA [23] provide a thorough assessment of language however these measures have been reported to be too lengthy for use in certain clinical contexts [2]. While these tests report published psychometric data, surprisingly, a systematic review examining these tests’ psychometrics was unable to identify published diagnostic (sensitivity/specificity) validation data verifying these tests’ abilities to identify aphasia in stroke populations [14].

The Brisbane EBLT aims to provide a time-efficient, comprehensive language test which assesses language across the severity spectrum in all language domains, is adaptable to varying clinical contexts and presents quick, evidence-based guidance regarding the need for intervention. As diagnostic analysis has been completed for all five tests, each version acts as a standalone assessment capable of identifying aphasia at the patient’s hospital bedside or clinic room in as little as 15–25 min. Adapted test scores further accommodate the needs of the patient and clinical context, enabling diagnostic estimates to be calculated despite the presence of fatigue or common co-occurring post-stroke conditions. This test aims to contribute to the accuracy and evidence-base of aphasia assessment procedures.

### Strengths and limitations

A strength of this validation study is in the methodology used. Study participants were a consecutive sample which met *a priori* sample size calculation requirements. The study employed a

Table 3. Patient sample demographics.

Characteristic	Patient sample demographic (n = 100)
Age	66.49 years (SD 15.06) (range 26–97 years)
Education	11.58 years (SD 3.149) (school and tertiary formal education only)
Sex	Males 57% (57) Females 43% (43)
Language	Monolingual (English) 84% (84) Bi or multilingual 16% (16)
Handedness	Right 86% (86) Left 8% (8) Ambidextrous 6% (6)
Average time between stroke onset and completion of both assessments	7 days 3.75 hours (range 1 day 21 h–28 days 6 h)
Stroke Type	Infarct 85% (85) <ul style="list-style-type: none"> <li>• 17 thrombolysis</li> <li>• 4 clot retrieval</li> <li>• 3 both clot retrieval and thrombolysis</li> </ul> Haemorrhagic 10% (10) Haemorrhagic and infarct 5% (5) <sup>a</sup>

<sup>a</sup>One patient had haemorrhagic transformation post lysis.

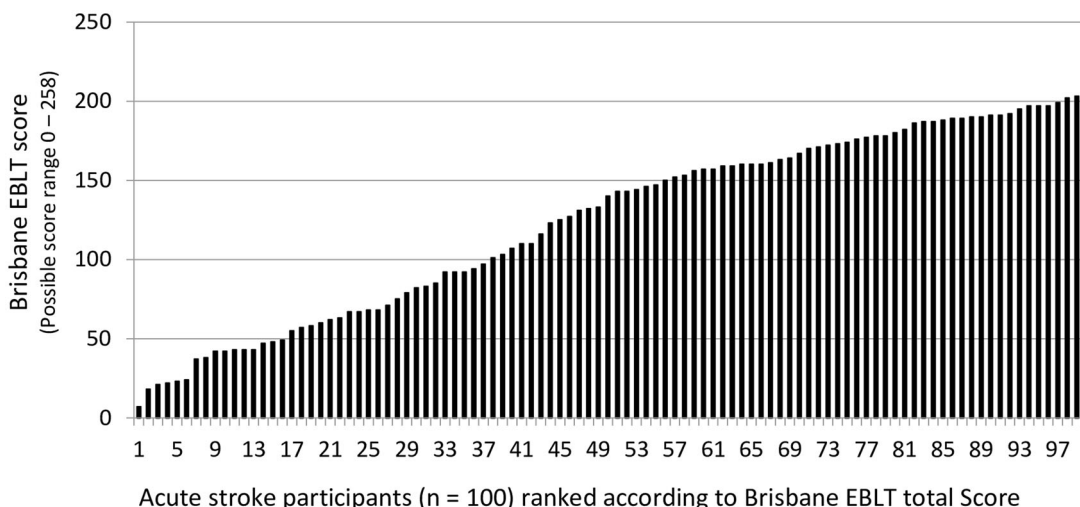


Figure 2. Distribution of severity of language impairment ranked by Brisbane EBLT score.

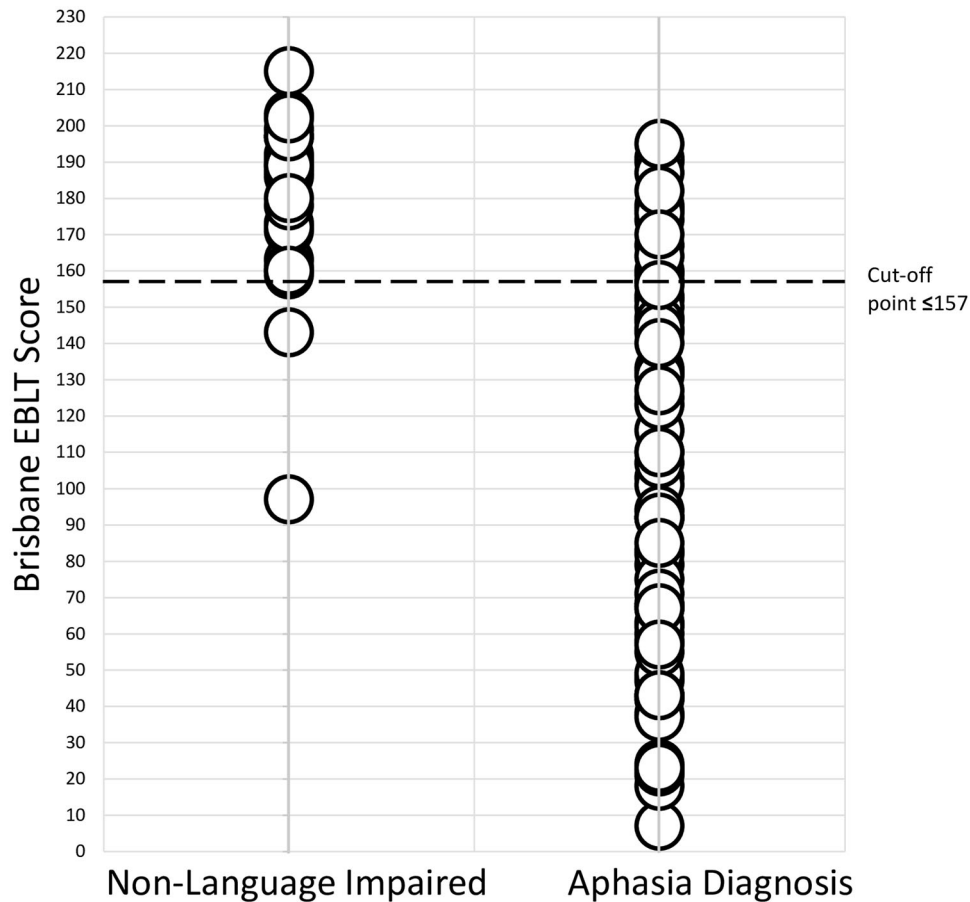


Figure 3. Brisbane EBLT score by language diagnosis.

robust reference standard which used multiple clinical factors to inform the diagnostic decision [21]. Finally, the index test and reference measure were randomized in order of administration and test administrators were blinded to the results of the other measure. This study aims to comply with recognised published STARD diagnostic reporting standards [18].

Different test versions optimise the clinical utility of this assessment. The shorter High Level Test and the full Brisbane EBLT reported identical diagnostic estimates. The High Level test consequently provides a short but psychometrically equivalent alternative to the longer comprehensive test and demonstrates that increased test length does not necessarily equate to increased diagnostic accuracy estimates. Conversely, the full 49 subtest Brisbane EBLT presents with no floor or ceiling effects and may be useful in a research or clinical context where patient performance of all ability levels can be compared across the same language measure. Adapted test scores, which enable the test to be adjusted for co-occurring conditions (e.g., hemianopia, apraxia of speech) provide additional flexibility. The presence of co-morbidities was common within the recruited sample with hemianopia, hemiparesis, hearing impairment, upper limb apraxia, apraxia of speech and dysarthria all present within the patient cohort. The Brisbane EBLT was specifically designed to minimise the impact of such co-occurring conditions (e.g., large text/stimulus items, not penalising for motor based impairment in responses and allowing repetition of tasks to account for deficits such as hearing impairment). As a result, less than 5% of patient language responses were unable to be scored thus indicating the new test's ability to adapt to the complex and multifaceted needs of stroke populations.

#### Statistical uncertainty

This study's results need to be interpreted in the context of the following factors. While both the High Level and full Brisbane EBLT report the same high specificity of 92.6%, their sensitivity is 80.8% indicating the test is less accurate in excluding a language disorder. However, comparison of aphasia patients who achieved Brisbane EBLT score of  $\leq 157$  and non-aphasia study participants indicated similar performance across the two groups. Hence, despite these lower diagnostic estimates, any patients not identified through this cut-off will have a language test performance highly similar to the language-intact stroke group. The clinical impact of any possible language impairment upon patient functioning is likely to be minimal. However, given the clinical usefulness of being able to exclude impairment, higher cut-offs have also been provided which prioritise sensitivity  $\geq 90\%$ . Patients scoring above these cut-offs have low probability of aphasia. Finally, given the Foundation Tests' focus on severe language deficits, the resultant lower sensitivities of these measures indicate other versions of the assessment are more suited to identifying milder language deficits. All five Brisbane EBLT tests reported AUC results of  $\geq 0.808$  indicating strong overall discriminative performance.

#### Generalisability

This study is an initial examination of the diagnostic accuracy of the Brisbane EBLT however the generalisability and reproducibility of these results within practice or in similar study populations have not yet been verified. A separate validation study in a

Table 4. Brisbane EBLT diagnostic accuracy characteristics.

Brisbane EBLT test version	Possible score range	Patient scores	Cut-off score (≥90% specificity)	Sensitivity (95% CI)	Specificity (95% CI)	LR+	LR-	At risk zone (≥90% sensitivity)	Cut-off score (≥90% sensitivity)	Sensitivity (95% CI)	Specificity (95% CI)	LR+	LR-	Area Under Curve, SE (95% CI)
Brisbane EBLT <sup>a</sup>	0–258	7–215	≤157	80.8% (69.9–89.1)	92.6% (75.7–99.1)	10.9	0.21	158–177	≤177 (≥178)	91.78% (83.0–96.9)	62.96% (42.4–80.6)	2.47	0.13	0.908, 0.030 (0.85–0.97)
Brisbane EBLT <sup>b</sup>	0–254	7–211	≤154	80.8% (69.9–89.1)	92.6% (75.7–99.1)	10.9	0.21	155–174	≤174 (≥175)	93.15% (84.7–97.7)	59.26% (38.8–77.6)	2.28	0.11	0.907, 0.030 (0.85–0.97)
Brisbane EBLT <sup>c</sup>	0–105	6–101	≤89	82.19% (71.5–90.2)	85.19% (66.3–95.8)	5.54	0.21	90–94	≤94 (≥95)	90.41% (81.8–95.6)	40.74% (20.3–65.2)	1.52	0.23	0.862, 0.038 (0.79–0.94)
<b>Short Tests</b>														
Foundation Test <sup>a</sup>	0–67	7–67	≤61	57.5% (45.4–69)	96.3% (81–99.1)	5.53	0.44	Alternative test versions should be used to exclude condition						
Foundation Test <sup>b</sup>	0–64	7–64	≤60	68.5% (56.6–78.9)	96.3% (81–99.1)	18.49	0.33	Alternative test versions should be used to exclude condition						
Foundation Test <sup>c</sup>	0–55	7–55	≤51	68.5% (56.6–78.9)	96.3% (81–99.1)	18.49	0.33	Alternative test versions should be used to exclude condition						
Foundation Test <sup>d</sup>	0–52	7–52	≤48	65.75% (53.2–75.7)	96.3% (81–99.1)	17.75	0.36	Alternative test versions should be used to exclude condition						
Standard Test <sup>a</sup>	0–169	7–141	≤90	72.6% (60.9–82.4)	92.6% (75.7–99.1)	2.47	0.13	91–111	≤111 (≥112)	91.78% (83.0–96.9)	62.96% (42.4–80.6)	2.47	0.13	0.907, 0.030 (0.85–0.97)
Standard Test <sup>b</sup>	0–165	7–135	≤86	72.6% (60.9–82.4)	92.6% (75.7–99.1)	2.47	0.13	87–107	≤107 (≥108)	91.78% (83.0–96.9)	62.96% (42.4–80.6)	2.47	0.13	0.907, 0.030 (0.85–0.97)
Standard Test <sup>d</sup>	0–137	5–110	≤61	67.1% (55.3–77.3)	92.6% (75.7–99.1)	9.06	0.36	62–80	≤80 (≥81)	93.15% (84.7–97.7)	59.26% (38.8–77.6)	2.28	0.11	0.906, 0.030 (0.85–0.97)
High Level Test	0–176	0–136	≤78	80.8% (69.9–89.1)	92.6% (75.7–99.1)	2.75	0.12	79–96	≤96 (≥97)	91.78% (83.0–96.9)	66.67% (46.0–83.5)	2.75	0.12	0.909, 0.030 (0.85–0.97)

<sup>a</sup>Requires everyday objects. Test items require two of each of the following everyday objects (cup, spoon, pen and knife).  
<sup>b</sup>Adapted score: excluding hospital ward items. Enables the test to be administered in a clinic or non-hospital ward location. Total score is calculated by omitting the subtest “Naming objects from the room” which requires ward objects to be present (bed, pillow etc).  
<sup>c</sup>Adapted score: excluding verbal expression subtests. Enables the test to be administered with patients with co-occurring expressive deficits affecting their verbal responses (e.g., apraxia of speech). Total score is calculated by omitting the Verbal Expression section (subtests 15–29) & Delayed Recall (subtest 39) and summing performance from the remaining four modalities (auditory comprehension, perception, reading and writing).  
<sup>d</sup>Adapted score: excluding reading/writing. Enables the test to be administered with patients with significant fatigue or co-occurring deficits affecting reading (e.g., hemianopia) and/or written responses (e.g., hemiparesis). Total score is calculated by omitting reading and writing sections. Note, any dysgraphia and dyslexia will not be identified.

similar, independent population would demonstrate the test’s capabilities in comparable settings and enable a diagnostic systematic review and meta-analyses to be performed.

**Conclusion**

The Brisbane EBLT aims to provide a new language test to assist in the identification of aphasia. The test examines patient functioning across the severity spectrum in all language domains with Short Tests capable of identifying aphasia in just 15–25 min. As diagnostic estimates have been calculated for all test versions, each can act as a stand-alone assessment with different versions able to adjust to the needs of varying clinical environments and enable clinicians to select tests at their own discretion based on patient ability and the clinical context. These findings aim to improve the evidence-base of aphasia assessment procedures and assist with informing accurate aphasia epidemiological statistics, healthcare services, and developers of stroke guidelines. The Brisbane EBLT is available for download from [brisbanetest.org](http://brisbanetest.org).

**Acknowledgements**

ANZ Medical Trustees, Australian Stroke Foundation, Equity Trustees Wealth Services Ltd.; Royal Brisbane and Women’s Hospital, and Royal Brisbane and Women’s Hospital Foundation. The stroke units and speech pathology departments of the Royal Brisbane and Women’s Hospital and Princess Alexandra Hospitals in Brisbane made this study possible. The authors thank the stroke patients, clinicians and other study participants who contributed to this research. Full list of acknowledgements at [brisbanetest.org](http://brisbanetest.org).

**Disclosure statement**

EG is on the Clinical Council of the Stroke Foundation. No other conflicts of interest are reported.

**Funding**

ANZ Medical Trustees, Australian Stroke Foundation, Equity Trustees Wealth Services Ltd., Royal Brisbane and Women’s Hospital, and Royal Brisbane and Women’s Hospital Foundation. The funders had no role in design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

**ORCID**

Linda Worrall  <http://orcid.org/0000-0002-3283-7038>  
 Robyn O’Halloran  <http://orcid.org/0000-0002-2772-2164>

**References**

- [1] Hersh D, Wood P, Armstrong A. Informal aphasia assessment, interaction and the development of the therapeutic relationship in the early period after stroke. *Aphasiology*. 2018;32:876–901.
- [2] Vogel AP, Maruff P, Morgan AT. Evaluation of communication assessment practices during the acute stages post stroke. *J Eval Clin Pract*. 2010;16:1183–1188.
- [3] Kertesz A. *Western Aphasia Battery – Revised*. San Antonio (TX): Pearson; 2007.

- [4] Goodwin D. Homonymous hemianopia: challenges and solutions. *Clin Ophthalmol*. 2014;8:1919–1927.
- [5] Simpson B, McCluskey A, Lannin N, et al. Feasibility of a home-based program to improve handwriting after stroke: a pilot study. *Disabil Rehabil*. 2016;38:673–682.
- [6] Enderby PM, Wood VA, Wade DT, et al. The Frenchay Aphasia Screening Test: a short, simple test for aphasia appropriate for non-specialists. *Int Rehabil Med*. 1987;8:166–170.
- [7] Flamand-Roze C, Falissard B, Roze E, et al. Validation of a new language screening tool for patients with acute stroke: The Language Screening Test (LAST). *Stroke*. 2011;42:1224–1229.
- [8] Doesborgh SJC, van de Sandt-Koenderman WME, Dippel DWJ, et al. Linguistic deficits in the acute phase of stroke. *J Neurol*. 2003;250:977–982.
- [9] Howard D, Swinburn K, Porter G. *Comprehensive Aphasia Test*. New York (NY): Psychology Press; 2004.
- [10] Goodglass H. *The Assessment of Aphasia and Related Disorders*. 3rd ed. Philadelphia (PA): Lippincott Williams & Wilkins; 2001.
- [11] LaPointe L. *Handbook of aphasia and brain-based cognitive-language disorders*. New York (NY): Thieme Medical Publishers; 2011.
- [12] Crary MA, Haak NJ, Malinsky AE. Preliminary psychometric evaluation of an Acute Aphasia Screening Protocol. *Aphasiology*. 1989;3:611–618.
- [13] El Hachoui H, Visch-Brink EG, de Lau LML, et al. Screening tests for aphasia in patients with stroke: a systematic review. *J Neurol*. 2017;264:211–220.
- [14] Rohde A, Worrall L, Godecke E, et al. Diagnosis of aphasia in stroke populations: a systematic review of language tests. *PLoS One*. 2018;13:e0194143.
- [15] Straus S, Glasziou P, Richardson S, et al. *Evidence based medicine. How to practice and teach it*. 4th ed. London (UK): Churchill Livingstone; 2010.
- [16] Buderer NM. Statistical methodology: I. Incorporating the prevalence of disease into the sample size calculation for sensitivity and specificity. *Acad Emerg Med*. 1996;3:895–900.
- [17] Scarpa M, Colombo A, Sorgato P, et al. The incidence of aphasia and global aphasia in left brain-damaged patients. *Cortex*. 1987;23:331–336.
- [18] Bossuyt PM, Reitsma JB, Bruns DE, et al.; STARD Group. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ*. 2015;351:h5527.
- [19] Binder JR, Frost JA, Hammeke TA, et al. Human brain language areas identified by functional magnetic resonance imaging. *J Neurosci*. 1997;17:353–362.
- [20] Yorkston KM, Beukelman DR. An analysis of connected speech samples of aphasic and normal speakers. *J Speech Hear Disord*. 1980;45:27–36.
- [21] Reitsma JB, Rutjes AW, Khan KS, et al. A review of solutions for diagnostic accuracy studies with an imperfect or missing reference standard. *J Clin Epidemiol*. 2009;62:797–806.
- [22] Whurr R. *Aphasia screening test: a multi-dimensional assessment procedure for adults with acquired aphasia*. Milton Keynes (UK): Routledge; 2011.
- [23] Ellmo W, Graser J, Krchnavek B, et al. *Measure of Cognitive-Linguistic Abilities (MCLA)*. Norcross (GA): The Speech Bin; 1995.
- [24] Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res*. 1999;8:3–15.
- [25] Wilson SM, Eriksson DK, Schneck SM, et al. A quick aphasia battery for efficient, reliable, and multidimensional assessment of language function. *PLoS One*. 2018;13:e0192773.