Cultural adaptation and psychometric testing of The Scenario Test UK for people with aphasia

Katerina Hilari[1], Lara Galante[*1], Anneline Huck[1], Madeleine Pritchard[1], Lucy Allen[2], Lucy Dipper[1]

[1]Division of Language and Communication Sciences, School of Health Sciences, City, University of London

[3] Guy's and St Thomas' NHS Foundation Trust

Corresponding author

Professor Katerina Hilari

Division of Language and Communication Science

School of Health Sciences

City, University of London

Northampton Square

London EC1V 0HB

UK

k.hilari@city.ac.uk

Tel: +44 (0) 207 040 4660

How to cite

Hilari, K., Galante, L., Huck, A., Pritchard, M., Allen, L. and Dipper, L. (2018), Cultural adaptation and psychometric testing of The Scenario Test UK for people with aphasia. International Journal of Language & Communication Disorders. doi:10.1111/1460-6984.12379

To access on publisher's website:

http://onlinelibrary.wiley.com/doi/10.1111/1460-6984.12379/full

To access at City, University of London's OpenAccess repository, go to

http://openaccess.city.ac.uk/view/creators/Hilari=3AK=2E=3A=3A.html

Key words: aphasia, functional communication, instrument development, psychometrics, outcome

[*] Lara Galante sadly passed away in 2015.

Declaration of Interest

Part of this study was funded by a doctoral studentship awarded to [second author name], by [name of University, name of School] Research Sustainability Fund.

Abstract

Background*:* This study explores the psychometric properties of The Scenario Test UK, a culturally adapted version of the Dutch original, which evaluates functional, daily-life communication in aphasia. The Scenario Test assesses communication in an interactive context with a supportive communication partner.

Aims*:* To evaluate the reliability (internal consistency, inter-rater and test-retest reliability) and construct validity (convergent, discriminant and known-groups validity) of The Scenario Test UK.

Methods and Procedures*:* The Scenario Test UK and other language, cognition and praxis assessments were administered to persons with aphasia after stroke (3+ months post stroke) and to non-aphasic controls. Participants were recruited primarily through community stroke groups. Measures were completed in an interview format. Standard psychometric criteria were used to evaluate reliability and construct validity.

Outcomes and Results: 74 participants with aphasia and 20 participants without aphasia took part. The Scenario Test UK showed high levels of reliability. Internal consistency (Cronbach's α = 0.92), inter-rater reliability (ICC = 0.95) and test–retest reliability (ICC = 0.96) were excellent. Inter-rater agreement in scores on the individual items ranged from good – excellent (κ = 0.41–1.00) for all but two items (4c κ = 0.38, 6c κ = 0.36). The Scenario Test UK demonstrated good levels of convergent (ρ = 0.37–0.75) and discriminant validity (ρ = -.04–.23). There was strong evidence for known groups validity (U = 132.50, $p < .001$), with those with aphasia scoring significantly lower [median (IQR) = 47 (39.8–51)] than those without aphasia [53 (52–54)].

Conclusions and Implications: The data support the reliability and validity of the Scenario Test UK as an assessment of functional, daily-life communication for persons with aphasia. Further testing is needed in independent samples on the measure's psychometric properties, including its sensitivity to change. Pending this testing, The Scenario Test UK can be used as an assessment tool to evaluate communication skills with people with aphasia, to guide goal setting for therapy and to measure outcomes in response to therapy.

What this paper adds to existing knowledge

What is already known on this subject

- A thorough assessment of aphasia requires the assessment of functional communication, which provides an understanding of how aphasia impacts a person's daily life.
- The Scenario Test, developed in Dutch by van der Meulen *et al*. (2010) is a test that measures functional communication taking a multi-modal approach, making it appropriate for individuals with limited verbal production.

What this study adds

- This paper describes the adaption of The Scenario Test for use in the UK (The Scenario Test UK) and provides evidence of its reliability and construct validity.

Clinical Implications

- Psychometric testing supports the reliability and validity of the Scenario Test UK. It is a promising new measure for the assessment of functional, daily-life communication for people with aphasia.

Introduction

Aphasia can have a profound effect on functional communication, that is, the ability to communicate effectively in everyday activities and situations (Armstrong, Ferguson, and Simmons-Mackie 2013). It reflects a person's ability to understand or to convey a message independent of the chosen modality, in order to communicate successfully (Frattali 1992). The ability to communicate in daily life settings is highly important, and affects a person with aphasia's ability to maintain relationships (Northcott, Marshall, and Hilari 2016), social participation (Northcott, Moss, Harrison, and Hilari 2016), and health-related quality of life (Fotiadou, Northcott, Chatzidaki, and Hilari, 2014; Hilari, Needle, and Harrison 2012). While assessment of aphasia has traditionally focused on the language impairment, there is now growing consensus that a thorough assessment needs to take a more holistic approach.

In the UK, the *Royal College of Speech and Language Therapists Clinical Guidelines* (2005) recommend that assessment comprises: (i) language impairment, (ii) functional communication, and (iii) psychological well-being. Within the framework of the International Classification of Functioning, Disability and Health (ICF) (World Health Organization (WHO), 2001), functional communication relates to the levels of activity and participation. An impairment of functional communication could, for example, lead to a difficulty in making a GP appointment or in conversing with friends, which refers to activities. If impairment of functional communication leads to more long-term effects, such as losing friends (Northcott and Hilari 2011), it relates to the ICF level of participation and to psychological well-being. Frattali and colleagues argue that assessments of functional communication need to target the activity level of the ICF model in order to be valid (Frattali, Thompson, Holland, Wohl, and Ferketic 1995). Recent studies have also investigated what goals people with aphasia have in relation to the ICF framework, and further, what outcomes are most important to people with aphasia and their families (Wallace *et al*. 2016; Worrall *et al*. 2011).

The majority of goals that people identified as a priority were related to the levels of activity and participation. This reflects the relevance of everyday activities (Worrall *et al*. 2011) and demonstrates that an assessment of functional communication is important to people with aphasia.

Despite the importance of functional communication, there is a lack of appropriate diagnostic tools to measure it, particularly those that are appropriate for individuals who have severe aphasia, and limited verbal production.

Functional communication assessment

While assessments of language impairment provide a detailed picture of the severity and type of aphasia including spontaneous speech, comprehension, repetition, naming, reading and writing, (e.g., the *Frenchay Aphasia Screening Test* (FAST), Enderby, Wood, Wade, and Hewer 1987; the *Comprehensive Aphasia Test* (CAT), Swinburn, Porter, and Howard 2004), assessments of functional communication aim to capture the effect that a language impairment has on the ability to communicate in natural contexts. This can allow the generation of treatment goals that reflect natural communication (Hartley 1990).

An assessment of functional communication may reveal similar results to an assessment of language (Bakheit, Carrington, Griffiths, and Searle 2005; Frattali *et al*. 1995; Irwin, Wertz, and Avent 2002; Laska, Bartfai, Hellblom, Murray, and Kahan 2007; Meier, Johnson, Villard, and Kiran 2017), but may also diverge. Lomas *et al*. (1989) found that the *Communicative Effectiveness Index* (CETI) did not correlate with outcomes of the *Western Aphasia Battery* (WAB) (Kertesz 1982). This outcome is not surprising, given that everyday communication settings offer extra-linguistic cues that can have a facilitative role. A functional approach to communication assessment takes account of these extra-linguistic factors as well as non-linguistic ways of communicating (Hartley 1990). For example, in addition to spoken language, a person with aphasia may point to objects, make gestures, write a message down, or draw it. Using all available communicative means is referred to as total

communication (Rautakoski 2011). If people use these alternative and additional means to communicate, communication may be much more successful than indicated by a score on a language assessment alone. Studies looking at whether change in language measures is associated with change in functional communication measures following therapy have found that though both types of measures may find improvements following therapy, language change scores are not significantly correlated with functional communication change scores (Aftonomos, Steele, Appelbaum, and Harris 2001; Meier *et al*. 2017). In other words, *patterns of improvement* at the impairment and functional levels also *diverge* and therefore both types of assessment should be used regularly.

Many of the earlier assessments of functional communication are rating tools based on observation of communication, such as the *Functional Communication Profile* (FCP) (Sarno 1965), the *Communicative Effectiveness Index* (CETI) (Lomas *et al*. 1989), and the *American Speech and Hearing Association Functional Assessment of Communication Skills for Adults* (ASHA-FACS) (Frattali *et al*. 1995). These tools involve a person familiar to the person with aphasia (PWA) judging their ability to communicate effectively. The ASHA-FACS, for example, provides a list of statements on daily life activities such as using a TV and a radio, and the therapist or a significant other is asked to rate how well the PWA is able to perform the task described in each statement.

Whereas most assessments focus on verbal communication, the *Multimodal Communication Screening Test for Persons with Aphasia* (MCST-A) (Lasker and Garrett 2006) allows the PWA to use modalities other than speaking (e.g., gestures, pantomime, or air-writing) in multiple different communicative tasks. Individuals are also classified as to whether they are dependent or independent of a communication partner. The *Assessment of Communicative Effectiveness in Severe Aphasia* (ACESA) (Cunningham, Farrow, Davies, and Lincoln 1995) is so far the only assessment that focuses on individuals with severe aphasia. It measures the ability to communicate in semi-structured conversations, and includes object and picture descriptions.

The *Communicative Abilities of Daily Living Test* (CADL and CADL-2) (Holland 1980; Holland, Frattali, and Fromm 1999) and the *Amsterdam-Nijmegen Everyday Language Test* (ANELT) (Blomert, Koster, and Kean 1995) assess functional communication in daily life situations. Like the MCST-A, the CADL-2 encourages the use of total communication. Individuals receive a variety of communicative tasks, some of which involve verbal responses, but many others can be solved by pointing to the correct stimulus picture or by manipulating concrete objects (e.g. using coins to "pay"). The ANELT, which measures functional communication in role-play settings, is a tool appropriate mostly for individuals with mild aphasia. The PWA has to imagine being in a particular scenario, and is asked to convey a message while the experimenter acts as an involved listener, but does not engage in the communication (Ruiter, Kolk, Rietveld, and Lotgering 2011). Both the comprehensibility and the intelligibility of the message are assessed.

In summary, a range of functional communication assessments is available. Some have very good psychometric properties (e.g., ANELT, ASHA-FACS, CADL-2), suggesting they are strong measures. However, the assessments described thus far have limitations. Some of the batteries are subjective, involving *rating* communicative skills rather than *measuring* them directly (FCP, CETI, ASHA-FACS). Some tools measure communication but are not focused on daily life communication (ACESA, MCST-A), whilst others do not capture the role of the communication partner (CADL), or focus primarily on individuals with mild aphasia (ANELT). These issues motivated the development of *The Scenario Test*.

The Scenario Test

The Scenario Test is a Dutch functional communication assessment (van der Meulen, van de Sandt-Koenderman, Duivenvoorden, and Ribbers 2010). This test measures how PWA convey a message (verbally and/or non-verbally) in daily-life situations; examines communication in an interactive setting; and is designed for use with individuals with severe aphasia. So far, the Scenario Test is

available in Dutch and is in the process of being adapted for German (Krzok and Plum 2016). The Scenario Test is based on the ANELT and maintains the use of daily-life situations, but adds two new facets: it allows multi-modal communication, and creates an interactive setting which mirrors everyday situations. In a series of scenarios presented in black and white drawings the person with aphasia is asked to adopt the role of a character who is faced with a communicative task. In a shopping scenario, for example, they are looking to buy a sweater and the shop assistant asks: "Can I help you?". The PWA can convey the message by speaking or writing, or by non-verbal communication such as gestures, drawing, or using a communication device. If the response is not correct or not clear, the examiner acts as a facilitative communication partner, and moves through a series of prompts. Scores can range 0–54 with higher scores indicating better functional communication. The Scenario Test also elicits information on the type (speaking or writing) and frequency (sometimes, often, only) of verbal communication, frequency (sometimes, often, only) and effectiveness (not, sometimes, mostly) of non-verbal communication (gesture, drawing, device), the flexibility in shifting the communicative mode (never, some after help, some spontaneous, often after help, often spontaneous), the quantity (none, sometimes, a lot) and type of help needed from the examiner, and the comprehension of the scenarios (poor, reasonable, good).

The original Scenario Test was administered to 122 people with aphasia, of whom 43 were unable to communicate verbally, and to 25 control participants. The analysis demonstrated high internal consistency (Cronbach's $\alpha = 0.96$; item-rest correlations = 0.58–0.82), test-retest reliability (ICC = 0.98), and inter- and intra-rater reliability (ICC = 0.86–1.00). The measure also had good convergent validity, as demonstrated by moderate to strong (0.50–0.85) correlations with three other measures of communication in aphasia, and good known-groups validity, as established by group differences between PWA and controls, and between people unable to use speech and those communicating verbally. The Scenario Test was also sensitive to changes in performance over six months (van der Meulen *et al*. 2010).

Aims

The Scenario Test was adapted for use in the UK (The Scenario Test UK) and this study aimed to evaluate the reliability (internal consistency, inter-rater and test–retest reliability) and construct validity (convergent, discriminant and known-groups validity) of The Scenario Test UK against standard psychometric criteria.

Methods

The Scenario Test UK

The Dutch developers provided us with the original test, the test manual, and an English translation of the test; and they trained us in the administration of the test. Adapting the original Scenario Test into a version that could be tested in the UK involved 1) translating the manual (including administration instructions and acceptable responses), and checking it with the Dutch developers; and 2) commissioning a set of conceptually equivalent illustrations to convey the scenarios that were culturally appropriate for the UK. New images were commissioned to feature characters of more varied race and gender, and to update clothing and backgrounds to look modern. The images were created with less detailed backgrounds than the originals, as detailed backgrounds can distract from the main features of the concept to be conveyed. Finally, a new version of the main character was drawn with an attempt to create more ambiguity around its age and gender, to help examinees identify with it. The images of the UK version were reviewed by the Dutch developers to confirm conceptual equivalence. For more information on The Scenario Test UK and a sample image, see online appendix.

A small pilot study (n = 9) was conducted to check the acceptability of the UK adaptation, based on burden, completion rates and score distributions. Participants were recruited through the University's aphasia research register and comprised three post-stroke participants without aphasia and six participants with aphasia. All participants were able to complete the measure in an average

of 20 minutes, with no particular difficulties. As expected, they obtained a broad range of total scores: 27–54, with a mean (SD) of 47.0 (10.5).

Design

A cross-sectional interview-based survey study was carried out to objectively evaluate the psychometric properties of The Scenario Test UK.

Participants

Participants were recruited through stroke community groups and the University's aphasia research register. We aimed to recruit at least 60 participants with aphasia and 20 participants with stroke without aphasia. Participants were eligible to take part if they were $\geq$ 18 years old, had a stroke at least 3 months before taking part, and spoke fluent English prior to the stroke. Those with aphasia had to meet the same criteria and also have aphasia due to the stroke. Aphasia was screened with the FAST. We aimed to validate The Scenario Test UK in a stroke and aphasia population, therefore participants were excluded if they had severe cognitive impairment and uncorrected hearing or visual problems. This was to ensure that any language or functional communication problems participants had were due to their stroke and aphasia rather than cognitive impairment or sensory problems. Participants were also excluded if they had terminal health conditions. Cognitive impairment was screened with the cognitive subtest of the CAT, which mainly looks at semantic and short-term memory. Participants who performed under 50% on either of these subtests were not included. Hearing and visual problems were determined through observation and self-report and health condition was determined by self-reports and/or reports of a friend or relative.

Procedure and Measures

Ethical approval was obtained from City, University of London's School of Health Sciences (the LCS Proportionate Review Committee). Participants were seen either at home or at the University Clinic,

by one of four administrators. Three of the four administrators were Speech and Language Therapists with experience of working with people with aphasia and the fourth was a clinical linguist. All administrators received training on the administration of The Scenario Test UK, which comprised an introduction to the measure, watching five to six videos of the test being administered to different people with aphasia; administering the test to at least two volunteers with aphasia and receiving feedback; and scoring three to four videos independently and comparing scores. Participants were given aphasia-friendly written information about the project, which the administrator went through with them. Written consent was obtained at least 48 hours after the project was first explained to them so that participants could make an informed decision. Assessments were completed in two to three sessions of approximately 1–1 ½ hours. An extra session of about 20 minutes was administered to 20 consenting participants 7–14 days later to evaluate the test-retest reliability of The Scenario Test UK. These participants completing the extra session were recruited by asking each participant as they came into the study, until the required number was reached.

Measures included the FAST and CAT cognitive screen as indicated above, The Scenario Test UK, the *Cognitive Linguistic Quick Test* (CLQT) (Helm-Estabrooks 2001) non-verbal subtests, the *Birmingham University Praxis Screen* (BUPS) (Bickerton, Riddoch, Samson, Balani, Mistry and Humphreys 2012), the *Limb Apraxia Screen* (LAS) (Poeck 1986), and the ASHA-FACS, which was completed by a partner or significant other of the person with aphasia. The FAST comprises four sections that assess both expressive and receptive language: speaking, writing, comprehension of spoken and written words. Scores on the FAST range 0–30 and high scores indicate milder aphasia. The CAT cognitive screen is the first part of the CAT, which aims to determine if cognitive deficits may affect the performance of a person with aphasia on the CAT language battery. It includes subtests on visual field / visual neglect deficits, semantic and non-verbal episodic memory, arithmetic problems, and ideomotor / ideational apraxia. Scores range 0-38 and higher scores are indicative of better cognitive function. The CLQT comprises five nonverbal tasks and five linguistic

tasks. Six composite cognitive domain scores are calculated from these ten tasks: attention, memory, executive functions, language, and visuospatial skills. However, since attention, memory and executive function composite scores include linguistic tasks (story retelling, generative naming), which in people with aphasia may be more affected by aphasia rather than cognition, only the visuospatial composite score was used in this study. Visuospatial scores range 0–105, with higher scores indicating higher cognitive skills. The BUPS, though a short assessment, covers a wide range of praxis abilities: gesture production, gesture recognition, gesture imitation, and multiple object use. Higher scores indicate better praxis skills and the maximum score is 12 for gesture production, 6 for gesture recognition, 12 for imitation, and 12 for multiple object use. As the BUPS has a limited number of items, it was substantiated by the LAS, which comprises 10 meaningful pantomime gestures, produced to verbal command with a maximum score of 10. Lastly, the ASHA-FACS is an other-rated test of a person with aphasia's functional communication. It comprises 44 items which cover four domains: social communication; communication of basic needs; reading, writing and number concepts; and daily planning. Scores range 1–7 and higher scores indicate more independence in functional communication.

Psychometric testing and data analysis

Reliability is concerned with the stability and consistency of a measure. It refers to its homogeneity and the extent to which it is free from random error. If a measure is consistently yielding the same results when administered repeatedly and no change has occurred then it is free of random error. For The Scenario Test UK, we tested internal consistency, (the extent to which items in the scale measure the same construct); test-retest reliability, (whether the test yields the same/ very similar scores when administered twice and no real change has occurred); and inter-rater reliability, (whether two raters scoring the same administration of the measure will give it the same scores). Validity refers to the extent to which an instrument measures what it purports to measure. Construct validity emphasizes the meaning of participants' responses to an instrument. It involves

comparing the instrument to external criteria and accumulating evidence, such as correlations with measures measuring the same construct (convergent validity), differences with measures measuring different constructs (discriminant validity) and differences among groups that should differ on the measure of the construct (known groups validity).

For test-retest reliability, as indicated above, 20 participants completed The Scenario Test UK a second time 7–14 days after the first assessment. Three of the four administrators were involved as raters for inter-rater reliability: 25 of the original 74 videos of participants with aphasia (33%) were rescored by a different administrator/rater. In testing the convergent validity of The Scenario Test UK, we hypothesised that for people with aphasia the measure will correlate moderately-highly with measures of language and communication (FAST, ASHA-FACS). As indicated above (e.g., Frattali *et al*. 1995; Irwin, Wertz, and Avent 2002), language and functional communication measures typically correlate and the FAST correlates highly with the Functional Communication Profile (Sarno 1969). The ASHA-FACS, as a functional communication measure should also correlate with The Scenario Test UK, though moderate correlations may be expected as the ASHA-FACS is rated by others rather than directly measuring functional communication. We also hypothesised that The Scenario Test UK would correlate moderately with measures of cognition (CLQT visuospatial skills composite, CAT cognitive screen), as visuospatial skills are important in alternative modes of communication such as use of images and drawing; and cognitive skills are essential for functional communication (Frattali *et al*. 1995; Meier *et al*. 2017). Lastly, there is a strong positive relation between limb praxis ability and spontaneous gesture ability (Borod, Fitzpatrick, Helm-Estabrooks and Goodglass 1989); we therefore hypothesised that The Scenario Test UK would correlate moderately with the BUPS gesture production total, and the LAS. In terms of discriminant validity, we hypothesised that The Scenario Test UK would have low correlations with less related subdomains of the chosen assessments, i.e., constructs that do not underpin self-initiation of verbal and non-verbal communication. We anticipated that it would have low correlations with the BUPS gesture recognition total, BUPS gesture imitation total, and BUPS

multiple object use score. For known groups validity, we hypothesised that those with aphasia would score lower on The Scenario Test UK that those without aphasia.

The following criteria were used for psychometric testing: for good internal consistency Cronbach's Alpha should be $\alpha > .70$ (Nunnally and Bernstein 1994); for good inter-rater reliability of the overall measure, intra-class coefficients should be ICC ≥ .80 (Streiner and Norman 2008). Cohen's kappa coefficient analysis was used to measure inter-rater reliability for each item (kappa scores). Kappa values of $\kappa \geq .75$ are considered an excellent level of agreement, $\kappa = .75–.41$ considered as a good level of agreement and $\kappa \leq .40$ is considered a poor level of agreement (Landis and Koch 1977). Intra-class correlation coefficients were also used to investigate test-retest reliability: for good test-retest reliability ICC ≥ .75 (Streiner and Norman 2008). Correlational analysis (Spearman's rho) was undertaken to test convergent and discriminant validity and a Mann-Whitney U test (as the data was skewed) compared The Scenario Test UK scores of those with aphasia versus those without.

Results

Participant characteristics

Ninety-four people with stroke took part in the project, 74 (78.7%) of whom had aphasia. They were 0.5–27 years post stroke, with a mean (SD) of 67.2 (57.8) months post stroke. Sixty-six (70.2%) had an ischaemic stroke. The majority were male 56 (59.6%), white 74 (78.7%), and married / had a partner 54 (57.4%). Table 1 details the participant characteristics for those with aphasia (n = 74) and those without aphasia (n = 20).

[table 1 about here]

Participant scores on measures

Table 2 details participants' scores on The Scenario Test UK and on the measures of language, cognition and praxis. On The Scenario Test UK, participants with aphasia scored a mean (SD) of 43.80

(10.53) and their scores ranged 9–54. Participants without aphasia scored higher with a mean (SD) of 52.85 (1.04) and had a much narrower range of scores, 51–54 (figure 1). In terms of the qualitative information provided by The Scenario Test UK, participants without aphasia communicated through spoken verbal communication (100%) and did not use writing (except for one), gesture, drawing or an alternative communication device. They understood well all the scenarios presented. In contrast, 47% of participants with aphasia communicated only through speech (28% often and 14% sometimes through speech) and 11% did not use speech at all. The majority did not use writing (82%). Thirty-two participants with aphasia (44%) used gesture, 33% used drawing, and 7% an alternative communication device. Those that did use an alternative communication strategy (gesture, drawing, device) were 75-96% effective in using these strategies, at least sometimes or mostly. The majority needed some help (68%) or a lot of help (21%) from their communication partner. Sixty participants (83.3%) demonstrated good understanding of the scenarios, with only one person demonstrating poor understanding (1.4%).

On all measures of language and cognition, participants without aphasia scored substantially higher than those with aphasia, close to ceiling and with a narrower range of scores. Of the 74 participants with aphasia, four had missing data on the LAS, two on the CLQT, two on the BUPS and one on its gesture production subsection (total numbers of participants completing each measure is indicated in table 2~~; their results are not included here~~. Participants without aphasia scored higher on BUPS measures of gesture production and gesture recognition and the LAS. Scores on BUPS gesture imitation and object use were similar between the two groups.

[table 2 about here]

[figure 1 about here]

Reliability

The Scenario Test UK demonstrated high internal consistency, $\alpha = .92$, with item-total correlations ranging .48–.77 (n = 94). It also demonstrated high test-retest reliability, ICC = .96, (n = 20). In terms

of inter-rater reliability, agreement between raters on the total score for The Scenario Test UK was excellent, ICC = .95, (n = 25 videos, 33% of total). Agreement in scoring for individual items was analysed using Cohen's kappa (κ) analysis. Kappa values ranged from .36–1.00. Table 3 details kappa (κ) values, levels of significance and the levels of agreement for each individual item. Agreement was excellent – good (range = 1–.41) for all but two items (88.9%). The two items that had a poor level of agreement were item 4c κ = .38 and item 6c κ = .36.

[table 3 about here]

Validity

In terms of convergent validity, as expected The Scenario Test UK correlated moderately - highly with measures of language, communication, cognition and praxis (ρ = .37 - .75, ps ≤ .001) (table 4). Discriminant validity was also good with The Scenario Test UK having low correlations with measures of gesture recognition, imitation and multiple object use (ρ = -.04 - .23, ps all ns) (table 5). Scores of those without aphasia [median (IQR) = 53 (52–54)] were significantly higher than those of participants with aphasia [47 (39.8–51)] (Mann-Whitney U = 139.5, p < .001), providing evidence of the known-groups validity of The Scenario Test UK.

[tables 4 and 5 about here]

Discussion

The Scenario Test UK assesses the functional communication of people with aphasia, in daily life situations. The assessment allows respondents to use total communication, and provides an experienced communication partner to evaluate communication skill in a supportive environment. The qualitative data generated by The Scenario Test UK showed that, overall, participants understood the scenarios in the test and all were able to complete the test. The qualitative data also

provided useful information on the frequency and effectiveness of the communication strategies respondents used. In contrast to those without aphasia, participants with aphasia often needed help from their communication partner and a substantial proportion used alternative communication strategies effectively. The quantitative results of this study fully support the reliability and validity of the UK adaptation of the test. All measures of reliability (internal consistency, inter-rater reliability, and test–retest reliability) were significantly high. Inter-rater agreement in scores on the individual items ranged from good to excellent (for all but two items), and levels of both convergent and discriminant validity were good. There was strong evidence for known-groups validity, with those with aphasia scoring significantly lower than those without aphasia.

The significantly high reliability ratings in the current study echo those found in the original study (van der Meulen *et al*. 2010). The data from both studies indicates that all items within The Scenario Test and The Scenario Test UK are consistently measuring the same underlying construct; that trained raters are able to achieve high levels of agreement; and that the test is a stable measure of functional communication over a period of 7–14 days.

Agreement between raters for each item was also analysed. Items 5b and 5c had excellent levels of agreement whereas items 4c and 6c had poor levels of agreement. All other items had a good level of agreement. The items with poor agreement were a sub-part of section 4, 'Visit' ('you'd like to invite your friend to your house for coffee next time'), and a sub-part of section 6, 'Restaurant' ('you'd like the waiter to bring a spoon for your soup'). The lack of agreement here is hard to explain. There was nothing about the scoring guidance for either item that was different to the items with good and excellent level of agreement, and so it is unlikely that scoring guidance was at the root of the disagreement. Equally there is no obvious difference in the complexity or interpretability of the expected responses that would explain the difference. Despite these two items, the overall inter-rater reliability of The Scenario Test UK was high, indicating the complete test reliably yields good agreement across raters. It should be noted that the three raters in this

study were trained, which helped with the agreement ratings. It is advisable for future users to receive training prior to administering or analysing The Scenario Test UK.

In terms of convergent validity, as expected, The Scenario Test UK showed moderate to strong correlations with other measures of functional communication, with language measures, and with measures of those aspects of non-verbal cognition and praxis that are required by the test. The high correlation with the FAST was expected as typically people with good language skills manage to get their message across effectively, whereas those with poorer language skills have more difficulties. Similar correlations have been reported between other language and functional communication measures (e.g., WAB, Kertesz 2006; and CETI, Bakheit *et al*. 2005). The Scenario Test UK had a higher correlation with the FAST than the ASHA-FACS, which may appear surprising, as the ASHA-FACS is a measure of functional communication. The moderate rather than strong correlation with the ASHA-FACS may result from the different scoring methods in the two functional communication measures. One difference is that, for the ASHA-FACS, another person has to evaluate the functional communication of the PWA, whereas The Scenario Test UK is a direct assessment. A second difference is that the ASHA-FACS assesses the PWA's independence in communication whereas The Scenario Test UK measures the PWA's ability to communicate a message in collaboration with a skilled communication partner.

The correlation between results from The Scenario Test UK and measures on cognitive tasks supports the assumption that an impairment in functional communication can be closely associated with an impairment in non-verbal cognition. Cognitive skills are essential for functional communication (Frattali *et al*. 1995; Meier *et al*. 2017). Visuospatial skills are important when people with aphasia attempt to use alternative means of communication such as drawing or pictorial supports. Moreover, it has been suggested that composite cognitive scores, such as those derived by the CLQT, are strongly associated with functional communication scores (Meier *et al*. 2017) and this was the case in our study. Further, the significant correlation between The Scenario Test UK and measures of praxis indicate a relationship between the ability to convey a message and the ability to

gesture. This result is not surprising given the known role that gestures have in natural communication. If a person can gesture successfully, this modality can be used as a substitute for speaking. However, gestures used for daily communication are different from gestures used in diagnostic tests of apraxia (Borod *et al.* 1989). Investigating the relationship between limb apraxia and the use of spontaneous gestures serving a communicative purpose in more detail, Borod *et al*. used a battery of tests on gestural communication, limb praxis, aphasia severity, auditory comprehension, and nonverbal intelligence. There was a significant correlation between the patients' performance on the praxis test and their performance in using gesture in a communicative setting; the more severe the limb apraxia the poorer the gestural communication skills. This finding is consistent with the significant correlation between The Scenario Test UK and measures of praxis.

The original study carried out in Dutch used different measures for assessing aphasia and functional communication (Amsterdam-Nijmegen Everyday Language Test, *Aachen Aphasia Test* (Graetz *et al*. 1991) and Communicative Effectiveness Index) and these measures were also positively correlated with The Scenario Test. Therefore, the findings from both studies support the validity of The Scenario Test as a measure of functional communication, and this study goes a step further by also highlighting the role of non-verbal cognition and praxis in task success.

The validity of the test is further supported by the discriminant validity findings in which there was low to moderate correlation with less related constructs (gesture recognition, imitation and multiple object use). Furthermore, as expected, known groups validity was strongly indicated, with those with aphasia scoring significantly lower on The Scenario Test UK than those without aphasia. The median score values between groups showed a difference of 6.5 points, arguably lower than expected. We interpret this finding positively, as it suggests that even those with severe expressive aphasia were able to convey their messages somewhat effectively when using a total communication approach, and in collaboration with a supportive communication partner. This is in line with the literature on the positive impact of a partner's communication skills on the communication of people with aphasia (Simmons-Mackie, Raymer and Cherney 2016).

Indeed, a benefit of The Scenario Test over other functional communication assessment tools is that it measures communication in a dialogue, in which the examiner acts as a helpful conversation partner who can support the speaker with aphasia by encouraging the use of non-verbal communication strategies (van der Meulen *et al*. 2010). While some individuals can communicate independently, others depend on their communication partner to get a message across, for example by prompts to switch from a verbal to a non-verbal communication channel (van der Meulen *et al*. 2010), and The Scenario Test allows this independence to be assessed in a structured and objective manner. As indicated above, a literature review by Simmons-Mackie and colleagues suggests that training communication partners to facilitate the person with aphasia can lead to beneficial results (Simmons-Mackie *et al*. 2016). The Scenario Test can inform the focus of such communication partner training, enabling clinicians to tailor it to the specific difficulties of individual clients with aphasia.

A major advantage of The Scenario Test is that it measures the ability to communicate in any modality, be it verbal, non-verbal, or a combination of the above. This renders the test an excellent tool to examine functional communication of people who have severe aphasia with no or very limited verbal language. While many tests demonstrate floor performance for people who have severe language difficulties (van der Meulen *et al*. 2010), The Scenario Test can discriminate participants who can convey a message from those that cannot, independent of the level of verbal communication. This total communication approach makes The Scenario Test a useful outcome measure to assess the efficacy of Alternative Augmentative Communication (AAC) therapy (van der Meulen *et al*. 2010), which is well suited for individuals with severe aphasia (Jacobs, Drew, Ogletree, and Pierce 2004). AAC therapy focuses on the communicative mode in relation to the individual's skills and communicative needs (van der Meulen *et al*. 2010). To measure the efficacy of AAC therapy, it is necessary to determine if the treatment leads to improvements in delivering a message in everyday communicative settings, which is the focus of the Scenario Test (van der Meulen *et al*. 2010). Additionally, outcomes from the Scenario Test can facilitate clinicians in planning bespoke

AAC therapy for clients and thereby allow for generating treatment goals which reflect natural communication (Hartley 1990). The test reveals the type and frequency of an alternative communication mode (gesture, drawing, use of a device), how easily somebody can shift from one mode to the other, and how often help is needed from the communication partner. Through this analysis, the researcher or clinician is informed about what type of communication may work most efficiently for the PWA.

As with any test of functional communication, the Scenario Test UK faces the challenge to construct tasks that are representative of daily-life communication. In an attempt to mimic a natural type of communicative event, The Scenario Test UK describes a series of situations in which the person with aphasia is presented with a communicative task. It could be argued that the scenario description, and explanation of the task by the tester, give away linguistic content that could be repeated by those people with aphasia who have intact repetition abilities. This would mean that the answer does not reflect word retrieval with a communicative intent but merely the ability to repeat. ~~However, we think that this is unlikely. T~~ We do not think this is very likely. First, in most of the scenarios, the linguistic content given does not include the expected response. For example, in the Shop example in the supplement, only the first of the three items includes the expected response ("would like to buy a sweater"), the other two do not (expected responses, e.g., "it's too big", "how much does it cost?"). Second, the final prompt by the tester is usually not very helpful with respect to the task. For example, in the first item of the Shop scenario, the tester asks "You are in a shop and would like to buy a sweater." The tester then removes the picture and says "The lady asks: Can I help you?" ~~In the restaurant scene, for example, the person is asked "You are having a drink with friends in a restaurant. You need to go to the toilet, but you do not know where it is. How do you ask the waiter?"~~ Intact repetition abilities ~~would not help~~ could be used here, but the person with aphasia would have to repeat something the tester said earlier and not the last thing they heard the tester say. In fact, during this study we have only observed repetition occasionally. The times it did occur, were mostly in the example scenario in which some people repeated the word

"train station" in combination with a facial expression of a question.

Another point to note around The Scenario Test UK itself is the acceptability of using AAC, which may be linked to cultural background, age, personality, length and severity of aphasia, and access to speech and language therapy. A review of AAC applications for aphasia suggested that using AAC strategies can be associated with lost hope for speech recovery, and further noted that AAC skills are often not used in daily situations (van de Sandt-Koenderman, 2004). According to the accommodation theory (e.g. Giles et al., 1973), people tend to accommodate to the preferred way of speaking of the people around them. For people with aphasia being assessed with The Scenario Test this would mean that they tend to use speaking over the use of gestures, writing or drawing, because the experimenter models speaking. With respect to this study, nearly half of the participants only used speaking as means of communication. Nevertheless, during the course of this study, participants with aphasia did not raise any difficulties accepting AAC and most used some AAC. Exploring the acceptability of using Alternative Augmentative Communication in more depth would have been outside of the scope of this paper, but would be a valuable area to investigate in the future.

Conclusions

The Scenario Test UK is a useful, valid and reliable clinical tool for the assessment of functional, daily-life communication. Its psychometric properties are highly promising, and are in line with previous psychometric findings reported in the original paper by van der Meulen *et al*. (2010). Its method of assessment contrasts with those functional communication assessment tools that are subjective, involving *rating* communicative skills rather than *measuring* them directly (FCP, CETI, ASHA-FACS); those tools that are not focused on daily life communication (ACESA, MCST-A); those tools that do not capture the role of the communication partner (CADL); or those that focus primarily on individuals with mild aphasia (ANELT).

As is common with new measures, further testing is needed in independent samples on the measure's psychometric properties, including its sensitivity to change. Pending this testing, The Scenario Test UK can be used as an assessment tool to evaluate communication skills with people with aphasia, to guide goal setting for therapy and to measure outcome in response to therapy, including multi-modal and total communication therapy.

References

AFTONOMOS, L., STEELE, R., APPELBAUM, J., and HARRIS, V., 2001, relationships between impairment-level assessments and functional-level assessments in aphasia: findings from LCC treatment programmes. *Aphasiology*, **15(10–11)**, 951–964.

ARMSTRONG, E., FERGUSON, A. and SIMMONS-MACKIE, N., 2013, Discourse and functional approaches to aphasia. In I. Papathanasiou, P. Coppens, and C. Potagas (Eds.), *Aphasia and related neurogenic communication disorders* (pp. 217–231). Burlington, MA: Jones & Bartlett Learning.

BAKHEIT, A.., CARRINGTON, S., GRIFFITHS, S. and SEARLE, K., 2005, High scores on the Western Aphasia Battery correlate with good functional communication skills (as measured with the Communicative Effectiveness Index) in aphasic stroke patients. *Disability and Rehabilitation*, **27(6)**, pp.287-291.

BICKERTON, W., RIDDOCH, M., SAMSON, D., BALANI, A., MISTRY, B., and HUMPHREYS, G., 2012, Systematic assessment of apraxia and functional predictions from the Birmingham Cognitive Screen (BCoS). *Journal of Neurology, Neurosurgery and Psychiatry,* **83**, 513–521.

BLOMERT, L., KOSTER, C., and KEAN, M., 1995, *Amsterdam– Nijmegen Everyday Language Test. Dutch Version.* Lisse: Swets & Zeitlinger.

BOROD, J., FITZPATRICK, P, HELM-ESTABROOKS, N., and GOODGLASS, H., 1989, The relationship between limb apraxia and the spontaneous use of communicative gesture in aphasia. *Brain and Cognition,* **10**, 121-131.

CUNNINGHAM, R., FARROW, V., DAVIES, C. and LINCOLN, N., 1995, Reliability of the assessment of communicative effectiveness in severe aphasia. *European Journal of Disorders of Communication*, **30(1)**, 1–16.

FOTIADOU D., NORTHCOTT S., CHATZIDAKI A., and HILARI K., 2014, Aphasia blog talk: how does

stroke and aphasia affect a person's social relationships? *Aphasiology,* **28**, 1281-1300.

FRATTALI, C. (1992). Functional assessment of communication: merging public policy with clinical views. *Aphasiology*, **6(1)**, 63–83. http://doi.org/10.1080/02687039208248577

FRATTALI, C., THOMPSON, C., HOLLAND, A., WOHL, C. and FERKETIC, M., 1995, *Functional Assessment of Communication Skills for Adults*. Rockville, Md: American Speech and Hearing Association.

FRIDRIKSSON, C., NETTLES, M., DAVIS, M., MORROW, L., and MONTGOMERY, A., 2006, Functional communication and executive function in aphasia. *Clinical Linguistics and Phonetics,* **20(6)**, 401-410.

GRAETZ, P., DE BLESER, R. and WILLMES, K., 1991, *Aachen Aphasia Test. Dutch versio*n. Lisse: Swets & Zeitlinger.

HARTLEY, L., 1990, Assessment of functional communication. In D. Tupper and K. Cicerone (Eds.), *The Neuropsychology of Everyday Life: Assessment and Basic Competencies, Volume 2 of the series Foundations of Neuropsychology* (pp. 125–168). Springer US. http://doi.org/10.1007/978-1-4613-1503-2_6

HELM-ESTABROOKS, N., 2001, *Cognitive Linguistic Quick Test (CLQT)*. Oxford: Pearson Assessment.

HILARI K., NEEDLE J. and HARRISON K., 2012, What are the important factors in health-related quality of life for people with aphasia? A systematic review. *Archives of Physical Medicine and Rehabilitation,* **93 (S1)**, S83-S95. http://dx.doi.org/10.1016/j.apmr.2011.05.028

HOLLAND, A., 1980, *Communicative Abilities in Daily Living*. Baltimore: University Park Press.

HOLLAND, A., FRATTALI, C. M. and FROMM, D., 1999, *Communication activities of daily living (CADL-2*). Austin: Pro-Ed.

IRWIN, W.., WERTZ, R., and AVENT, J., 2002, Relationships among language impairment, functional communication, and pragmatic performance in aphasia. *Aphasiology*, **16**(8), 823–835.

JACOBS, B., DREW, R., OGLETREE, B. and PIERCE, K., 2004, Augmentative and Alternative Communication (AAC) for adults with severe aphasia: where we stand and how we can go

further. *Disability and Rehabilitation*, **26(21/22)**, 1231–1240. http://doi.org/10.1080/09638280412331280244

KERTESZ, A., 2006, *The Western Aphasia Battery*. New York: Grune & Stratton.

KRZOK, F. and PLUM, L., 2016, Szenariotest: Training zur Durchführung des neuen Testverfahrens (CF0208). Presentation at the 16. Jahrestagung der Gesellschaft für Aphasieforschung und -behandlung (GAB), Osnabrück, Germany.

LASKA, A., BARTFAI, A., HELLBLOM, A., MURRAY, V., and KAHAN, T., 2007, Clinical and prognostic properties of standardised and functional aphasia assessments. *Journal of Rehabilitation Medicine*; **39**: 387–392.

LASKER, J. and GARRETT, K., 2006, Using the Multimodal Communication Screening Test for Persons with Aphasia (MCST-A) to guide the selection of alternative communication strategies for people with aphasia. *Aphasiology*, ***20*(2/3/4)**, 217–232. http://doi.org/10.1080/02687030500473411

LOMAS, J., PICKARD, L., BESTER, S., ELBARD, H., FINLAYSON, A. and ZOGHAIB, C., 1989, The Communicative Effectiveness Index: development and psychometric evaluation of a functional communication measure for adult aphasia. *The Journal of Speech and Hearing Disorders*, **54(1)**, 113–124.

MEIER, E., JOHNSON, J., VILLARD, S., and KIRAN, S., 2017, Does naming therapy make ordering in a restaurant easier? Dynamics of co-occurring change in cognitive-linguistic and functional communication skills in aphasia. *American Journal of Speech-Language Pathology,* **26,** 266-280.

NORTHCOTT, S. and HILARI, K., 2011, Why do people lose their friends after a stroke? *International Journal of Language & Communication Disorders*, **46(5)**, 524–534. http://doi.org/10.1111/j.1460-6984.2011.00079.x

NORTHCOTT S., MARSHALL J. and HILARI K., 2016, What factors predict who will have a strong social network following a stroke? *Journal of Speech, Language and Hearing Research,* **59**, 772-783. doi:10.1044/2016_JSLHR-L-15-0201

NORTHCOTT S., MOSS B., HARRISON K. and HILARI K., 2016, A systematic review of the impact of stroke on social support and social networks: associated factors and patterns of change. *Clinical Rehabilitation,* **30(8)**, 811-831 doi: 10.1177/0269215515602136

OFFICE OF NATIONAL STATISTICS, 2010, Standard Occupational Classification 2010. Volume 3 The National Statistics Socio-economic Classification: (Rebased on the SOC2010) User Manual. Newport, South Wales: Palgrave McMillan. Accessed online 18/10/2017 http://webarchive.nationalarchives.gov.uk/20160105160709/http://www.ons.gov.uk/ons/guide-method/classifications/current-standard-classifications/soc2010/soc2010-volume-3-ns-sec--rebased-on-soc2010--user-manual/index.html

POECK, K., 1986, The clinical examination for motor apraxia*. Neuropsychologia,* **24**, 129–134.

RAUTAKOSKI, P., 2011, Training total communication. *Aphasiology*, **25(3)**, 344–365. http://doi.org/10.1080/02687038.2010.530671

ROYAL COLLEGE OF SPEECH and LANGUAGE THERAPISTS, 2005, *Royal College of Speech and Language Therapists Clinical Guidelines*. Oxon, UK: Speechmark Publishing.

RUITER, M., KOLK, H.., RIETVELD, T. and LOTGERING, E., 2011, Towards a quantitative measure of verbal effectiveness and efficiency in the Amsterdam-Nijmegen Everyday Language Test (ANELT). *Aphasiology*, **25(8)**, 961–975. http://doi.org/10.1080/02687038.2011.569892

SARNO, M., 1965, A measurement in functional communication in aphasia. *Archives of Physical Medication and Rehabilitation*, **47**, 50–56.

SIMMONS-MACKIE, N., RAYMER, A. and CHERNEY, L., 2016, Communication partner training in aphasia: an updated systematic review. *Archives of Physical Medicine and Rehabilitation*, **97(12)**, pp.2202-2221.

VAN DER MEULEN, A., HOUTHUIZEN, J. and VAN DE SANDT-KOENDERMAN, W., 2008, Meten van communicatie bij ernstige afasie De Scenario Test: een nieuwe test voor verbale én nonverbale communicatie bij afasiepatiënten. *Logopedie En Foniatrie*, **(4)**, 112–116.

VAN DER MEULEN, I., VAN DE SANDT-KOENDERMAN, W., DUIVENVOORDEN, H. and RIBBERS, G.,

2010, Measuring verbal and non-verbal communication in aphasia: reliability, validity, and sensitivity to change of the Scenario Test. *International Journal of Language & Communication Disorders / Royal College of Speech and Language Therapists*, **45(4)**, 424–35. http://doi.org/10.3109/13682820903111952

WALLACE, S., WORRALL, L., ROSE, T., LE DORZE, G., CRUICE, M., ISAKSEN, J., KONG, A., SIMMONS-MACKIE, N., SCARINCI, N., and GAUVREAU, C. A., 2016, Which outcomes are most important to people with aphasia and their families? An international nominal group technique study framed within the ICF. *Disability and Rehabilitation*, **39(14)**, 1364-1379. http://doi.org/10.1080/09638288.2016.1194899

WORLD HEALTH ORGANIZATION., 2001, Health Organization. International classification of functioning, disability and health (ICF). Geneva.

WORRALL, L., SHERRATT, S., ROGERS, P., HOWE, T., FERGUSON, A. and DAVIDSON, B., 2011, What people with aphasia want: their goals according to the ICF. *Aphasiology*, **25(3)**, 309–322. http://doi.org/10.1080/02687038.2010.508530

Table 1: Participant characteristics for people with aphasia and people with stroke without aphasia

Characteristic	Participants with aphasia n = 74	Participants without aphasia n = 20
Gender		
Male	44 (59.5%)	12 (60%)
Female	30 (40.5)	8 (40%)
Age		
Mean (SD)	60.9 (12.4)	50.8 (12.7)
Range	24–97	27–71
Ethnic group		
White	60 (81.1%)	14 (70%)
Asian	6 (8.1%)	4 (20%)
Black	7 (9.5%)	1 (5%)
Mixed /other	1 (1.4%)	1 (5%)
Marital status		
Single	19 (25.7%)	5 (25%)
Has partner	11 (14.9%)	2 (10%)
Married	33 (44.6%)	8 (40%)
Divorced or widowed	11 (14.9%)	5 (25%)
Socioeconomic status[1]	*n = 73*	
Higher managerial, administrative and professionals	22 (30.1)	3 (15%)
Intermediate occupations	15 (20.5%)	6 (30%)
Routine and manual occupations	36 (49.3)	11 (55%)
Stroke type		
Ischaemic	55 (74.3%)	11 (55%)
Haemorrhagic	9 (12.2%)	8 (40%)
Mixed	2 (2.7%)	1 (5%)
Unknown	8 (10.8%)	0
Time post onset (months)		

Characteristic	Participants with aphasia n = 74	Participants without aphasia n = 20
Mean (SD)	63.3 (50.6)	81.5 (79.1)
Range	8–289	6–324
Comorbid conditions		
None	55 (74.3%)	16 (80%)
One	18 (24.3%)	4 (20%)
Two	1 (1.4%)	0

[1]**:** Based on the UK Standard Occupation Classification 2010, of the Office of National Statistics (2010).

Table 2: Scores on The Scenario Test UK and other measures for participants with aphasia and participants with stroke without aphasia

Measure	Participants with aphasia (n = 74)	Participants with stroke without aphasia (n = 20)
The Scenario Test UK		
Mean (SD)	43.80 (10.53)	52.85 (1.04)
Median (IQR)*	47.00 (39.80–51.00)	
Min–Max	9.00–54.00	51–54
FAST		
Mean (SD)	17.00 (7.38)	29.00 (1.76)
Median (IQR)		30.00 (28.00–30.00)
Min–Max	3.00–29.00	23.00–30.00
ASHA-FACS		
Mean (SD)	5.62 (1.13)	N/A
Median (IQR)	5.80 (4.89–6.53)	
Min–Max	2.00–7.00	
CAT Cognitive Score		
Mean (SD)	19.20 (3.67)	30.30 (9.51)
Median (IQR)	19.50 (18.00–20.00)	
Min–Max	11.00–37.00	16.00–38.00
CLQT Non-verbal Composite Score	*n = 72*	
Mean (SD)	38.82 (13.80)	50.55 (9.03)
Min–Max	9.00–58.00	29.00–61.00
BUPS Gesture Production	*n = 71*	
Mean (SD)	9.54 (2.61)	11.70 (0.80)

Measure	**Participants with aphasia (n = 74)**	**Participants with stroke without aphasia (n = 20)**
Median (IQR)		12.00 (12.00–12.00)
Min–Max	2.00–12.00	9.00–12.00
BUPS Gesture Recognition	*n = 72*	
Mean (SD)	4.86 (1.48)	5.65 (0.93)
Median (IQR)		6.00 (6.00–6.00)
Min–Max	0.00–11.00	3.00–6.00
BUPS Gesture Imitation	*n = 72*	
Mean (SD)	13.93 (5.73)	13.70 (9.44)
Min–Max	6.00–24.00	0.00–24.00
BUPS Multiple Object Use	*n = 72*	
Mean (SD)	11.50 (1.59)	11.85 (0.67)
Median (IQR)	12.00 (12.00–12.00)	12.00 (12.00–12.00)
Min–Max	0.00–12.00	9.00–12.00
LAS	*n = 70*	
Mean (SD)	7.30 (2.39)	9.35 (0.81)
Min–Max	0.00–10.00	8.00–10.00

***** Median (IQR): Reported only for skewed variables (skewness > ±1)
FAST: Frenchay Aphasia Screening Test
ASHA-FACS: American Speech-Hearing Association Functional Assessment of Communication Skills for Adults
CAT: Comprehensive Aphasia Test
CLQT: Cognitive Linguistic Quick Test
BUPS: Birmingham University Praxis Screen
LAS: Limb Apraxia Screen

Table 3: Level of agreement (kappa value) for each question of The Scenario Test UK

Item	Kappa value (κ)	Significance Level	Level of Agreement
1a	.73	*p < .001*	Good
1b	.41	*p = 001*	Good
1c	.41	*p = .002*	Good
2a	.54	*p < .001*	Good
2b	.47	*p < .001*	Good
2c	.49	*p < .001*	Good
3a	.56	*p < .001*	Good
3b	.64	*p < .001*	Good
3c	.66	*p < .001*	Good
4a	.46	*p = .002*	Good
4b	.49	*p < .001*	Good
4c	.38	*p = .005*	Poor
5a	.46	*p < .001*	Good
5b	1	*p < .001*	Excellent
5c	.80	*p < .001*	Excellent
6a	.57	*p < .001*	Good
6b	.66	*p = .001*	Good
6c	.36	*p = .002*	Poor

Table 4: Convergent validity of The Scenario Test UK for participants with aphasia

The Scenario Test UK	FAST	ASHA-FACS	CLQT visuospatial skills composite	CAT cognitive screen	BUPS gesture production	LAS
ρ	.75***	.50***	.46***	.37***	.37***	.44***
n	74	74	72	74	71	70

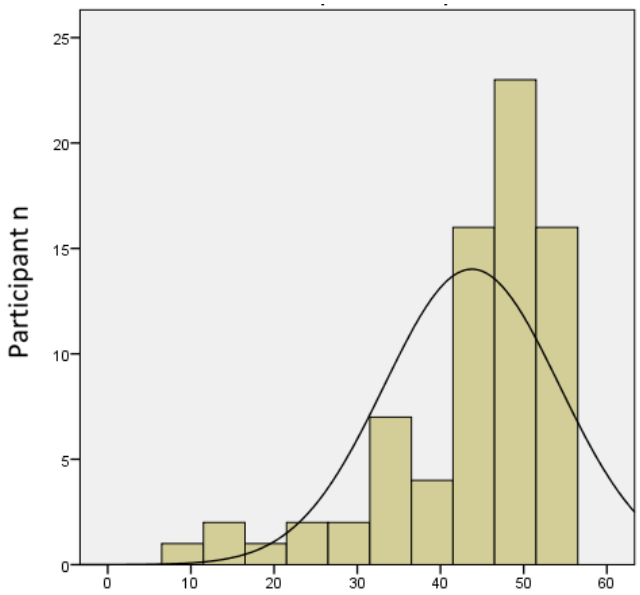
***: p ≤ .001
FAST: Frenchay Aphasia Screening Test
ASHA-FACS: American Speech-Hearing Association Functional Assessment of Communication Skills for Adults
CLQT: Cognitive Linguistic Quick Test
CAT: Comprehensive Aphasia Test
BUPS: Birmingham University Praxis Screen
LAS: Limb Apraxia Screen

Table 5: Discriminant validity of The Scenario Test UK for participants with aphasia

The Scenario Test UK	BUPS gesture recognition	BUPS gesture imitation	BUPS multiple object use
ρ	.23	.19	-.04
p	ns	ns	ns
n	72	72	72

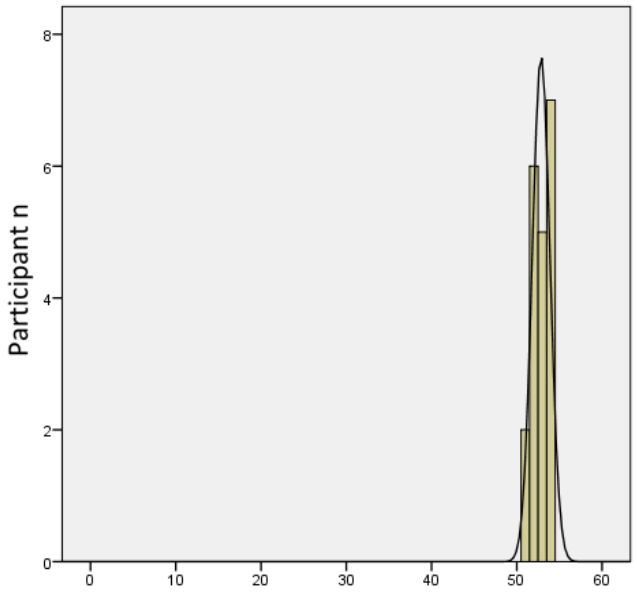
ns: non significant, p > .05.
BUPS: Birmingham University Praxis Screen

Figure 1: The Scenario Test UK scores' distribution of participants with aphasia and participants without aphasia



Participants with aphasia (n = 74)

Mean (SD) = 43.80 (10.53)



Participants without aphasia (n = 20)

Mean (SD) = 52.85 (1.04)