Received: 13 August 2022 | Accepted: 15 February 2023

DOI: 10.1111/1460-6984.12867

International Journal of Language & Communication Disorders

RESEARCH REPORT

Which blueberries are better value? The development and validation of the functional numeracy assessment for adults with aphasia

Kerri Ichikowitz | **Carolyn Bruce** | **Vanessa Meitanis** | **Kelly Cheung** | **Yekyung Kim** | **Esther Talbourdet** | **Caroline Newton**

Division of Psychology & Language Sciences, University College London, London, UK

Correspondence
Kerri Ichikowitz, Division of Psychology & Language Sciences, University College London, London, UK.
Email: k.ichikowitz.16@ucl.ac.uk

Funding information
This research did not receive any funding from agencies in the public, commercial or not-for-profit sectors.

Abstract

Background: People with aphasia (PWA) can experience functional numeracy difficulties, that is, problems understanding or using numbers in everyday life, which can have numerous negative impacts on their daily lives. There is growing interest in designing functional numeracy interventions for PWA; however, there are limited suitable assessments available to monitor the impact of these interventions. Existing functional numeracy assessments lack breadth and are not designed to be accessible for PWA, potentially confounding their performance. Additionally, they do not include real-life demands, such as time pressure, which may affect their ecological validity. Thus, there is a crucial need for a new assessment to facilitate further research of PWA's functional numeracy.

Aims: To develop, validate and pilot a wide-ranging, aphasia-friendly functional numeracy assessment to investigate how functional numeracy is impacted by aphasia severity and time pressure demands, and to explore predictors of PWA's functional numeracy.

Methods & Procedures: To develop the Functional Numeracy Assessment (FNA), 38 items inspired by the General Health Numeracy Test (GHNT) and Excellence Gateway were adapted for suitability for PWA and entered in a computerized psychometric-style test. The final 23 items (FNA23) were selected based on 213 neurotypical controls' performance, and controlled for difficulty, response modality and required numeracy skills. Aphasia-friendly adaptations of the GHNT and Subjective Numeracy Scale were used to examine the FNA23's concurrent validity. Internal consistency reliability and interrater reliability (for spoken responses) were also examined. A novel Time Pressure Task was created by slight adaptation of seven FNA23 questions to explore the effects of time pressure on functional numeracy performance. A total of 20 PWA and 102 controls completed all measures on an online testing platform.

Outcomes & Results: The FNA23 demonstrated acceptable internal consistency reliability (*KR-20* = 0.81) and perfect interrater reliability (for spoken responses). FNA23 and GHNT scores were positively associated, suggesting satisfactory concurrent validity. PWA demonstrated poorer functional numeracy than controls and took longer to complete assessments, indicating that aphasia impacts functional numeracy. Time pressure did not significantly impact performance. PWA demonstrated a wide range of functional numeracy abilities, with some performing similarly to controls.

Conclusions & Implications: The FNA23 is a wide-ranging, valid and reliable assessment which, with further development, will be a useful tool to identify and monitor PWA's functional numeracy difficulties in research and clinical practice. Considering PWA's widespread functional numeracy difficulties evidenced by this study, all PWA would likely benefit from routine evaluation for functional numeracy difficulties as part of their neurorehabilitation journeys.

KEYWORDS
acquired language disorder, aphasia, assessment, ecologically valid, functional numeracy, time pressure

WHAT THIS PAPER ADDS

What is already known on this subject

- Few studies have investigated functional numeracy difficulties in PWA. No published functional numeracy assessments exist that have been specifically designed to be accessible for PWA.

What this paper adds to existing knowledge

- The newly developed FNA23 is a valid and reliable tool to extensively assess PWA's functional numeracy. This study confirmed previous findings of widespread functional numeracy difficulties in PWA that are related to their aphasia severity.

What are the potential or actual clinical implications of this work?

- The FNA23 can be used to assess PWA's functional numeracy to inform areas of strengths and difficulties to target in intervention, and to monitor progress towards achieving intervention objectives. All PWA should be routinely evaluated for functional numeracy difficulties.

INTRODUCTION

Approximately 350,000 people in the UK have aphasia, a language disorder resulting from an acquired brain injury (Stroke Association, 2017). In addition to language impairments, people with aphasia (PWA) often have difficulties understanding or using numerical information in everyday life (Benn et al., 2022). Furthermore, PWA have identified improving their use of numbers in the context of communication and improving their ability to use and understand money when shopping as important treatment outcomes (Wallace et al., 2017). These everyday numeracy abilities are encompassed in the definition of *functional numeracy*: the knowledge and skills required to cope with the mathematical demands of everyday life contexts, enabling full participation in daily life activities (e.g., managing medication doses, choosing the best value for money groceries; Ginsburg et al., 2006). Consequently, difficulties with functional numeracy can have wide-ranging impacts on PWA's lives, for example, as societal barriers hindering their community participation (Howe et al., 2008) and negatively affecting their independence and quality of life (Benn et al., 2022).

Acquired numeracy difficulties in PWA, referred to as 'acalculia', have been a topic of research since as early as 1926 (Berger, as cited in Delazer et al., 1999). In their study of 50 participants with left hemisphere brain lesions, Delazer et al. (1999) found that participants who also had aphasia were most likely to present with transcoding and calculation difficulties, which correlated with their aphasia severity. This association between acalculia and aphasia severity has been replicated more recently (e.g., Gonzalez et al., 2021); however, findings from most studies have been based on traditional numeracy tests that assess these difficulties in isolation from the everyday life contexts in which they typically occur.

Given the accumulating reports of numeracy difficulties in clinical populations (Semenza, 2008), and the growing awareness of discrepancies in people's performance on traditional isolated tests compared with everyday life (Burgess et al., 2006), there has been emerging interest in developing ecologically valid 'function-led' numeracy assessments. These assessments enable more thorough investigation of functional numeracy than traditional numeracy tests, such as the arithmetic subtest in the Comprehensive Aphasia Test (Swinburn et al., 2004), can provide. One such assessment, developed for people with brain injuries, is the Ecological Assessment Battery for Numbers (EABN; Villain et al., 2015). Currently only available in French, the EABN, which contains 18 subtests covering eight themes (e.g., time, shopping), has satisfactory psychometric properties for people with brain injuries (based on 17 participants, mostly with left hemisphere lesions). Villain et al. (2015) highlighted that some participants took exceedingly long to complete the EABN, despite 'normal' scores (compared with 126 neurotypical controls). The authors suggested that these participants would be limited in everyday numeracy situations due to their slow execution of tasks, emphasizing the importance of considering participants' speed alongside their accuracy scores.

Another assessment of numerical abilities designed for clinical populations (specifically neurological diseases) is the Numerical Activities of Daily Living (NADL; Burgio et al., 2022; Semenza et al., 2014), which includes a participant and caregiver interview to assess participants' awareness of their difficulties, a 19-item informal test of functional numeracy and a formal test of traditional numeracy. Although the informal test covers various functional themes (e.g., time, transportation), the range of numeracy skills assessed is limited and approximately half the questions rely on participants' memory of numerical facts (e.g., current date, age, price of a new car).

Both the EABN and NADL have been used with PWA. Robert et al. (2021) assessed 36 French-speaking left hemisphere stroke patients and found that EABN scores correlated with aphasia severity (assessed by the Boston Diagnostic Aphasia Examination). Limited details are available in English about the EABN's format and delivery, preventing critical evaluation of these findings. Proios et al. (2021) administered the NADL to 33 PWA and concluded that PWA were likely to have difficulties with most numerical activities of daily living. However, the NADL's informal test was presented entirely verbally to participants, which may have confounded performance of PWA with auditory comprehension difficulties. It is a significant limitation that neither study discussed whether adaptations were made to improve the assessments' accessibility for PWA by making the format aphasia-friendly; for example, with simplified vocabulary and syntax, large text and emboldened key words to reduce the assessment's cognitive and comprehension load (Brennan et al., 2005; Rose et al., 2011). Therefore, findings based on the EABN and NADL must be interpreted cautiously as it is unclear whether they reflect a true relationship between aphasia and functional numeracy or an artefact of the assessments' language demands confounding PWA's performance (e.g., participants with more severe aphasia perhaps performed poorly because they had difficulties understanding the questions rather than difficulties performing the tasks).

Although aphasia-friendly assessments of functional numeracy skills do exist, they only exist as part of large assessment batteries, for example, the Communication Activities of the Daily Living (CADL-3; Holland & Fromm, 2018) and Functional Assessment of Communication Skills for adults (ASHA-FACS; Frattali et al., 1995). These measures lack the breadth required to assess PWA's functional numeracy comprehensively.

The development of a more extensive functional numeracy assessment suitable for PWA is much needed to facilitate further research of PWA's functional numeracy difficulties, inform intervention design, and monitor PWA's progress towards achieving intervention objectives, both in research and clinical practice. Therefore, the first aim of this study, which was not pre-registered, was to develop, validate and pilot a wide-ranging, aphasia-friendly assessment of functional numeracy skills: the 23-item Functional Numeracy Assessment (FNA23).

The FNA23 was designed to be self-paced, reflecting the design of existing functional numeracy assessments (e.g., NADL and EABN). However, given Villain et al.'s (2015) argument regarding participants response times, it was recognized that this traditional, self-paced format might limit the FNA23's ability to accurately reflect PWA's functioning in everyday situations, which often involve additional demands and complexities that may influence their performance, such as time pressure. Burgess et al. (2006) argued that these real-life demands must be considered in the design of assessments to increase ecological validity.

Time pressure has been suggested as a situational factor that affects neurotypicals' comprehension and interpretation of numerical information related to health decision-making (a subset of functional numeracy; Lipkus & Peters, 2009). Caviola et al. (2017) suggested that time pressure acts as a stressor leading to suboptimal strategy use and thus poorer mathematical performance, either by interfering with decision-making processes required to select the optimal strategy or by overloading working memory resources. Alternatively, Echt et al. (1998) hypothesized that poorer performance in tasks involving time pressure was due to increased demands on processing speed. Considering reports of both working memory difficulties (e.g., Laures-Gore et al., 2011) and slowed processing speed (e.g., Faroqi-Shah & Gehman, 2021) in PWA, it seems likely that PWA's functional numeracy performance would be affected even more than neurotypicals' in situations involving time pressure. However, no research has directly examined the effects of time pressure on PWA's functional numeracy. Therefore, the second aim of this study was to determine whether PWA's functional numeracy was affected by time pressure and thus whether the addition of time pressure to the FNA23's design might increase its sensitivity and ecological validity. It was hypothesized that PWA would perform more poorly on functional numeracy tasks when under time pressure compared with self-paced.

The third aim of the study was to investigate whether previous findings that aphasia impacts functional numeracy would be replicated when functional numeracy was measured using aphasia-friendly assessments with reduced language demands, reducing potential confounds on PWA's performance. It was hypothesized that PWA would score significantly more poorly and take significantly longer to complete functional numeracy tasks than neurotypical controls. It was also hypothesized that PWA would report a significant decline in their subjective functional numeracy abilities and preferences following their stroke.

Despite previous reports of a relationship between aphasia severity and functional numeracy, limited research has considered whether demographic factors may partly explain this relationship. In their study of 200 PWA, Gonzalez et al. (2021) found that, when writing ability was controlled for, education predicted PWA's performance on traditional numeracy tests, but not aphasia severity, age or time post-stroke. These findings were partly supported by Proios et al. (2021) who found that aphasia severity and education but not age were significant predictors of the NADL's informal test scores (time post-stroke was not considered in their analysis). These differing findings may reflect the inherent differences between traditional and functional numeracy assessments, but further research is needed before conclusions can be made. Therefore, this study's fourth and final aim was to investigate whether previous findings of a relationship between aphasia severity and functional numeracy could be replicated when functional numeracy was measured by aphasia-friendly assessments, and whether age, education and time post-stroke contributed to this relationship.

This pilot study therefore sought to answer the following questions:

- Is the newly developed FNA a valid and reliable assessment of PWA's functional numeracy in terms of concurrent validity, internal consistency reliability and interrater reliability?
- Does time pressure impact PWA's functional numeracy performance?
- Is functional numeracy impacted by aphasia?
- Is PWA's functional numeracy related to the severity of their aphasia, and does age, education and time post-stroke contribute to this relationship?

METHOD

Ethical considerations

The UCL Language and Cognition Departmental Ethics Committee approved this study (LCD-2019-03). PWA were emailed an aphasia-friendly poster and GDPR-compliant information and consent forms (Data Protection Act, 2018). Due to COVID-19, all research sessions were conducted online. PWA's informed consent was obtained at the start of each session: verbally on Zoom and electronically on Gorilla (an online experiment builder and testing platform; www.gorilla.sc). PWA were briefed on Zoom about the Gorilla tasks and told that they might find some tasks more challenging than others. PWA had opportunities to ask questions during initial Zoom sessions and were encouraged to contact the researchers if further questions arose. Controls received information sheets either via email or Prolific (online participant recruitment service; www.prolific.co) and provided informed consent on Gorilla.

Participants

A total of 25 PWA participated in the study (M_{age} = 61.1 years, SD = 10.32 years, range = 39.8–75.9 years; $M_{\text{time post-stroke}}$ = 8.54 years, SD = 5.12 years; 10 identified as females and 15 identified as males) (see Table 1 for demographics). A total of 20 of the PWA had mild aphasia according to the Western Aphasia Battery—Revised (WAB-R; Kertesz, 2007; $M_{\text{aphasia quotient}}$ = 81.07, SD = 14.02, range = 31–97) and 18 had fluent aphasia. PWA were recruited from aphasia charities, support networks and a university communication clinic's research register, with the opportunity to win one of six £25 vouchers. PWA's inclusion criteria were minimum 6 months post-stroke, aged 18 or older, self-reported normal or corrected-to-normal vision and hearing, fluent in English pre-stroke and considered to have aphasia based on published WAB-R Aphasia Quotient (AQ) cut-off scores (Kertesz, 2007). Participants who classified as 'normal' on the WAB-R but self-identified as having aphasia and were considered to have aphasia based on the researcher's clinical judgement (e.g., word-finding difficulties in discourse; Fromm et al., 2017) were also included. PWA were excluded if they did not have internet access and a smartphone, tablet, or computer or did not complete all sessions by the end of the data collection phase (minimum 3 weeks after session 1). Two PWA withdrew during session two (demographics included in Table 1 with consent), and three were excluded due to not completing the Gorilla sessions, so in total, 20 PWA completed the study. The PWA sample size was limited due to the research team's limited resources to collect data for this pilot study.

A total of 102 neurotypical adults provided control data for the functional numeracy measures (69 identified as females and 33 identified as males; M_{age} = 49.69 years, SD = 17.77 years, range = 18.58–86.33 years; age was not recorded for 15 participants). Controls were recruited via Prolific for paid participation (n = 21 in 30–55 age range, n = 54 in 59–95 age range), an opportunity sample and a university Psychology Subject Pool (for course credits). Inclusion criteria included UK nationals who were fluent in English (self-reported), had internet and computer access and were aged 18 and above (older than 30 for Prolific to broaden the controls' age range). Controls who took part in prior FNA pilot studies were excluded. Figures 1 and 2 illustrate distributions of PWA and controls' current occupation and highest level of education.

A Mann–Whitney U-test indicated that PWA's ages (median = 59.9 years) were significantly higher than controls (median = 56.7; U = 641, z = −2.04, n = 116, p = 0.036). Controls' age had a bimodal distribution. Excluding participants younger than 30 years old yielded a non-significant age difference between PWA (median = 59.9) and controls (median = 58.58; U = 641, z = −754, n = 95, p = 0.451). To identify potential confounding effects from the age difference, analyses were conducted with the full sample and repeated with participants below 30 years old excluded.

Materials

Demographics questionnaire

Self-reported socio-demographic information was collected for all participants, including age, gender, occupation (previous if retired), highest level of education, whether English was their primary language and other languages spoken. Additional information collected for PWA included time post-stroke and hearing status.

Western Aphasia Battery—Revised (WAB-R)

Four WAB-R language subtests (spontaneous speech, auditory comprehension, repetition and naming; Kertesz, 2007) were conducted with PWA to determine the presence, severity and type of aphasia. Lower AQ scores indicated greater aphasia severity. The assessment was adapted for remote Zoom delivery based on Dekhtyar et al. (2020). Major adaptations included presenting picture description and auditory word recognition stimuli via screen share and providing remote cursor control for participants to indicate answers for auditory word recognition. Where participants could not remotely control the cursor (e.g., on tablets), carers would report what they pointed to or PWA were shown numbered stimuli and said the number corresponding to their answer. To limit interference from difficulties reading or saying the correct number, the assessor confirmed participants' answers by circling their chosen item. For the object naming subtest, PWA were shown images of objects on screen share, rather than physical objects, to ensure consistency and avoid lighting difficulties. For sequential commands, PWA were asked to have a pen, book and comb with them and position their camera so the assessor could observe their movements.

TABLE 1 Demographic information of participants with aphasia

	Age (years)	Gender	English as the primary language	Additional languages	Time post-stroke (years)	Current occupation (previous if retired)	Highest level of education	AQ	Aphasia type (fluent/non-fluent)[a]
P1	59.5	Female	Yes		2	Retired (worked for a charity)	GCSE	85.4	Anomic (fluent)
P2	39.8	Female	Yes	Urdu, Punjabi	3	Project manager	UG degree	97.2 [b]	Anomic (fluent)
P3	71.8	Male	Yes		9	Retired (hay merchant)	GCSE	30.9	Global (non-fluent)
P4	56.3	Male	Yes		12	Retired (accountant)	Accountanc degree	73.5	Wernicke's (fluent)
P5	75.8	Female	No	Krio	14	Retired (micro-biologist)	UG degree	82.4	Anomic (fluent)
P6	68.7	Male	Yes		15	Retired (customer service director)	A-level	82.4	Anomic (fluent)
P7	69	Male	Yes		4	Retired (photographer and designer)	UG degree	95.1[b]	Anomic (fluent)
P8	75.9	Female	Yes	French (limited)	18	Retired (solicitor)	Solicitor	89.9	Anomic (fluent)
P9	43.1	Male	Yes	Russian, Polish, Irish	11	Retired (financial computing)	PG degree	81.1	Anomic (fluent)
P10	71.7	Male	Yes	French, Japanese	21	Retired (solicitor)	PG degree	78.9	Anomic (fluent)
P11	53.8	Male	Yes		2	Photo editor	UG degree	94.1[b]	Anomic (fluent)
P12	55.8	Female	No	Italian	7	Translator	UG degree	89.9	Anomic (fluent)
P13[d]	67.3	Female	Yes		4	Retired	GCSE	86	Conduction (fluent)
P14	59.9	Male	Yes		8	Retired (Rail Operations Planner)	A-level	89.7	Anomic (fluent)
P15	58.9	Male	Yes		8	Aphasia charity	UG degree	77.7	Anomic (fluent)
P16	73	Male	Yes	French	11	Retired (accountant)	CIPFA	79.3	Anomic (fluent)
P17	46.9	Female	Yes	Ghanaian	6	Looking for work	Higher National Diploma	87.1	Anomic (fluent)
P18	45	Male	Yes	Hindi	4	Looking for work	PG degree	92.8	Anomic (fluent)
P19[c]	59.3	Female				Retired (HR)		54.3	Wernicke's (fluent)
P20[c]	71.1	Female	Yes		3	Retired (photo-journalist and lecturer)	PG degree	37.9	Broca's (non-fluent)

(Continues)

International Journal of Language & Communication Disorders

TABLE 1 (Continued)

	Age (years)	Gender	English as the primary language	Additional languages	Time post-stroke (years)	Current occupation (previous if retired)	Highest level of education	AQ	Aphasia type (fluent/non-fluent)[a]
P21[d]	56.1	Male			8	Retired (telecommunications)		89.1	Anomic (fluent)
P22	63.1	Male	Yes		13	Retired (librarian)	PG degree	73.4	Transcortical motor (non-fluent)
P23	62	Female	Yes		11	Retired (solicitor)	PG degree	78.6	Anomic (fluent)
P24	61.7	Male	Yes		6	Retired (Upholsterer)	No qualifications	76.1	Conduction (fluent)
P25[d]		Male			5	(Forensic accountant)		91.3	Anomic (fluent)

Notes: UG, undergraduate; PG, postgraduate; AQ, Western Aphasia Battery—Revised aphasia quotient (Kertesz, 2007).
[a]Fluency determined by the WAB-R classification.
[b]Classified as anomic due to word finding difficulties in discourse (Fromm et al., 2017), despite scoring above WAB-R cut-off of 93.8.
[c]Dropped out.
[d]Excluded from analyses due to not completing all sessions.

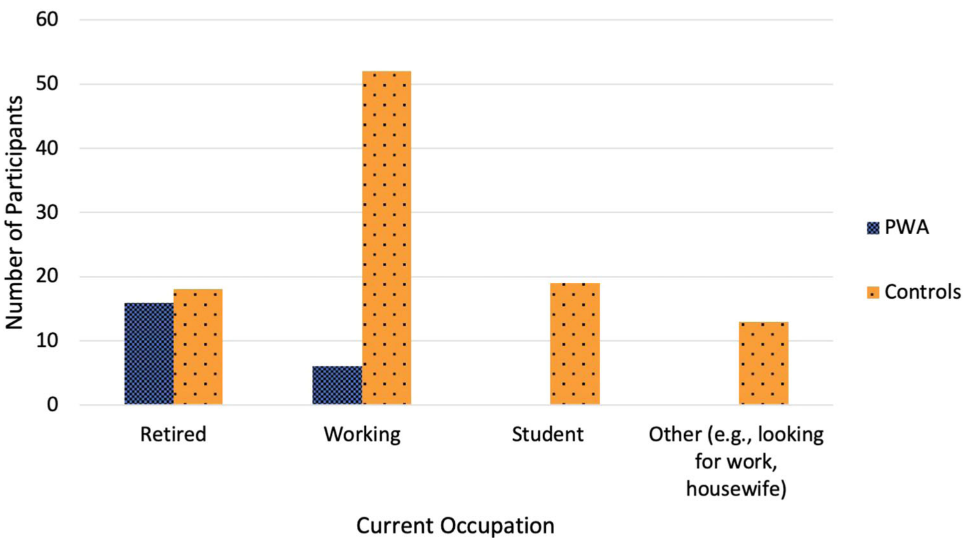

FIGURE 1 Participants' current occupation. [Colour figure can be viewed at wileyonlinelibrary.com]

Functional Numeracy Assessment (FNA)

The FNA, inspired by the General Health Numeracy Test (GHNT; Osborn et al., 2013) and Excellence Gateway (2020), was designed as a wide-ranging, objective assessment of PWA's functional numeracy skills. To make the assessment aphasia-friendly, language demands were minimized, and items included written questions, pictures, and automatically playing audio of the researcher saying the questions (to benefit participants with impaired reading comprehension; Rose et al., 2011). Because PWA have been found to have a range of strengths and weaknesses in different numeracy scenarios, questions varied in the predominant numeracy skills assessed (i.e., number identification, counting, computation, transcoding and number comparison), topic (e.g., travel and shopping), and response modality, including multiple-choice with typed or picture options, spoken (audio-recorded) and typed.

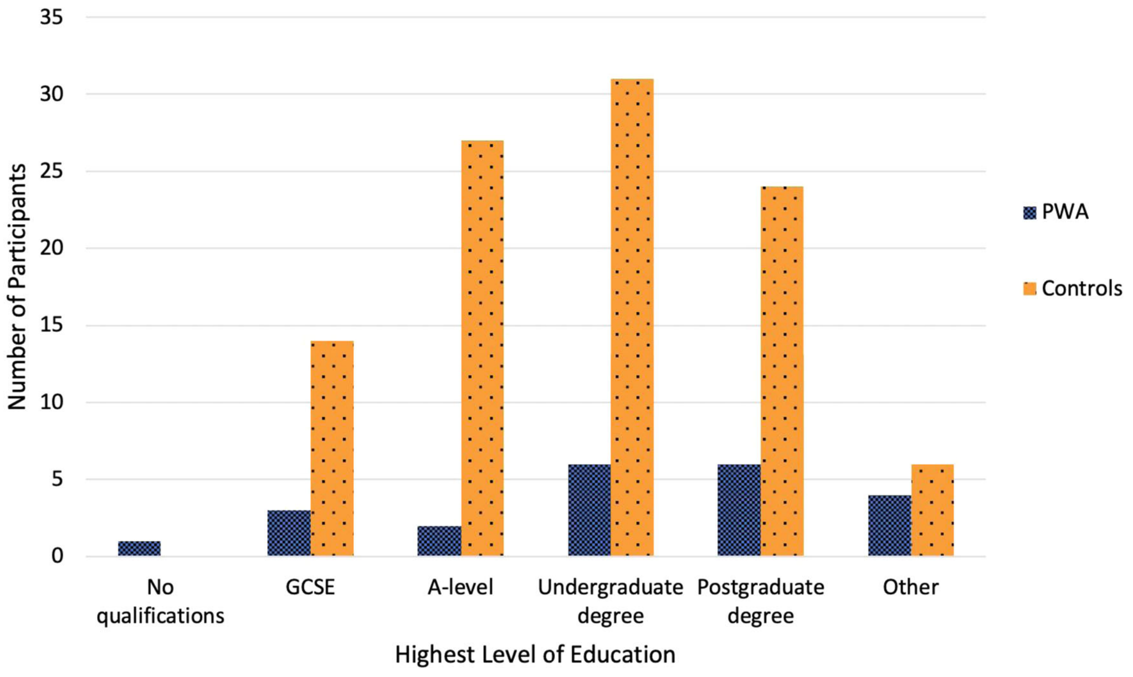

FIGURE 2 Participants' highest level of education. [Colour figure can be viewed at wileyonlinelibrary.com]

Each question's response modality was based on what the research team deemed most functional in that context (e.g., reading aloud a telephone number; see Figure 3 for examples and Appendix A for the questions).

Data were gathered on the initial 38-item version of the FNA from 213 neurotypical participants (M_{age} = 24.81 years, SD = 11.87 years, range = 18–81 years). Due to a technical fault, responses for six items with spoken response modalities were not saved and therefore could not be analysed. The assessment's internal consistency reliability was measured using the Kuder–Richardson coefficient (*KR-20* = 0.363). *KR-20* values generally range from 0.0 to 1.0, with values > 0.7 considered acceptable (Salkind, 2010). Item analysis was then used to refine the FNA. First, items with negative point-biserial correlations from an SPSS Reliability Analysis were removed (Pearson Education, 2015). To design a broad and varied assessment, the research team then considered the difficulty level (based on means), response modality and numeracy skill of items. Reliability analyses were rerun until a favourable combination was reached with the highest possible *KR-20* value (*KR-20* = 0.416), resulting in 17 items being selected. The technical fault with the six spoken items was fixed and these items were kept in the assessment so that the spoken response modality could be examined in the present study. This yielded the 23-item FNA (FNA23) which was used in this study.

Gorilla recorded how long participants took to complete each item and this was summed to calculate the time taken to complete the whole assessment. Gorilla scored most questions automatically, except spoken questions which two researchers scored. Questions were scored as correct or incorrect (self-corrections in spoken answers were considered correct), resulting in a total out of 23, with higher scores indicating stronger functional numeracy. Interrater reliability, evaluated by double scoring 10% of PWA and controls' spoken answers, showed 100% agreement.

Time Pressure Task

To assess functional numeracy skills under time pressure, the seven-item aphasia-friendly Time Pressure Task was designed by adapting items from the FNA23 (see Appendix B). Items chosen reflected a variety of skills and tasks typically performed under time pressure in everyday life (e.g., calculating change). To maximize the task's functionality, participants were first shown visual stimuli for 2 s, then they heard and saw the questions while a pulsing stopwatch indicated remaining time (see Figure 4 for an example). If time ran out before participants responded, the next item appeared automatically. Timings were based on pilot data from neurotypicals and aimed to add time pressure while still providing enough time to complete the question (varied per question between 20 and 35 s).

Although everyday numeracy scenarios do not always involve access to written prompts, Time Pressure Task questions included text to limit auditory processing difficulties confounding performance. The text was hidden after each question's audio finished playing, increasing the memory demands and similarity to functional scenarios.

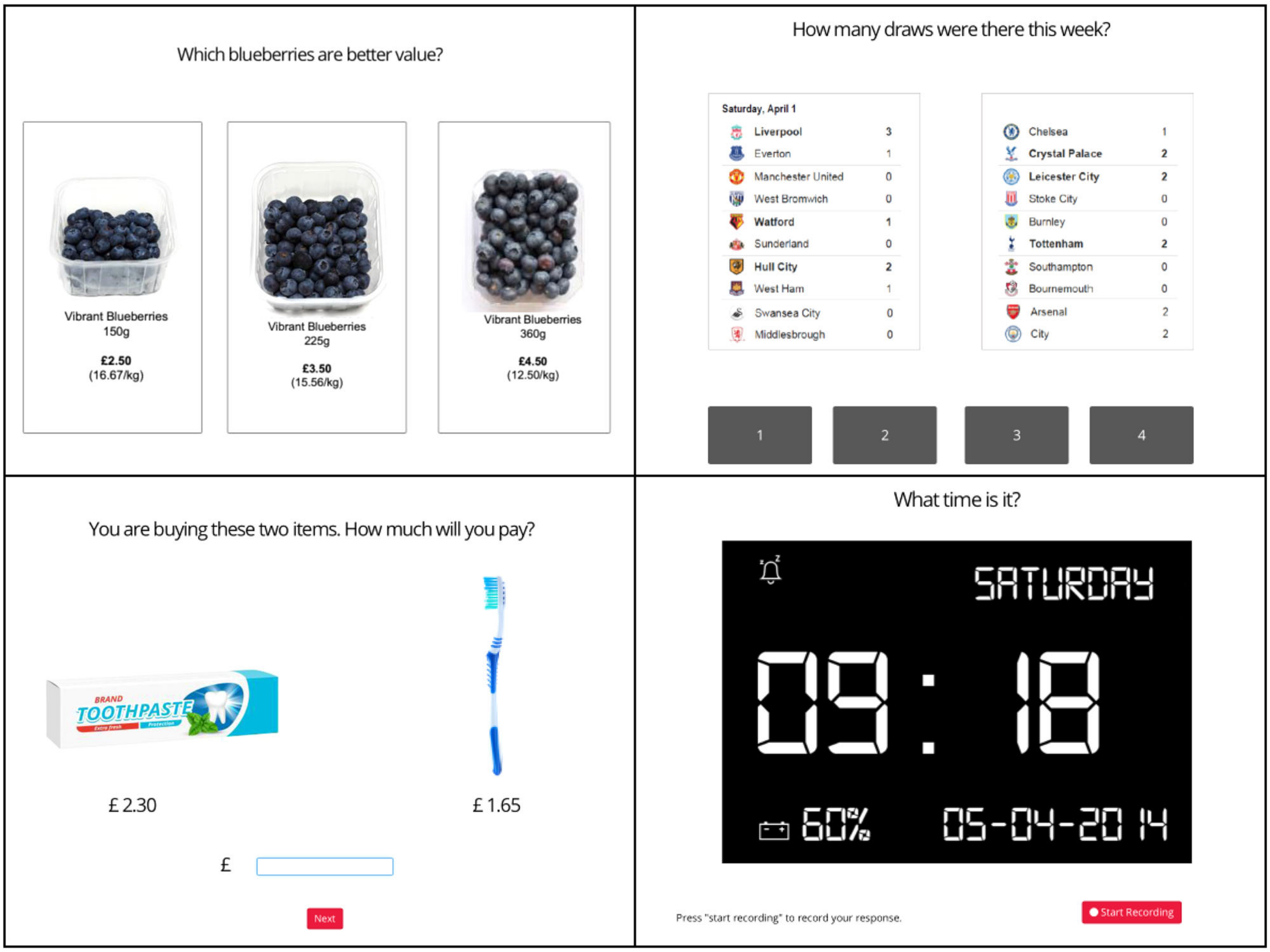

FIGURE 3 Screenshots of four questions in the FNA23 showing different response modalities. [Colour figure can be viewed at wileyonlinelibrary.com]

Questions were scored as correct or incorrect, resulting in a total out of seven, with higher scores indicating greater functional numeracy skills under time pressure. Participants' time to complete the task was also recorded. Additional scores out of six were calculated for matched questions from the Time Pressure Task (questions 1–6) and FNA23 (questions 7–9, 11, 15 and 17) so they could be directly compared.

Subjective Numeracy Scale (SNS)

The SNS (Fagerlin et al., 2007; Zikmund-Fisher et al., 2007) is a valid and reliable self-report questionnaire measuring participants' subjective ability to perform different numerical tasks and their preference for numerical information compared with language. It includes eight questions (four ability and four preference) answered on six-point Likert-type scales (question 7 is reverse coded). Fagerlin et al. (2007) validated the SNS with 287 neurotypical adults (Cronbach's alpha = 0.82) and reported a significant correlation ($r = 0.68$) with objective numeracy. Participants' time to complete the SNS was recorded and average ratings out of six were calculated across all eight questions, with higher ratings indicating stronger subjective functional numeracy.

The SNS was adapted for the present study to be more aphasia-friendly by bolding keywords, adding visuals and audio (Figure 5). The questions' wording was not altered to minimize affecting the assessment's psychometric properties.

The SNS's subjective nature enabled measurement of PWA's perceptions of premorbid functional numeracy through retrospective ratings. Therefore, PWA completed two SNS versions: First, the post-stroke SNS (identical to the version completed by controls), then the pre-stroke SNS, which asked about their abilities and preferences from before their stroke. Comparing ratings from the pre- and post-stroke SNS indicated PWA's self-perceived changes in functional numeracy following their stroke.

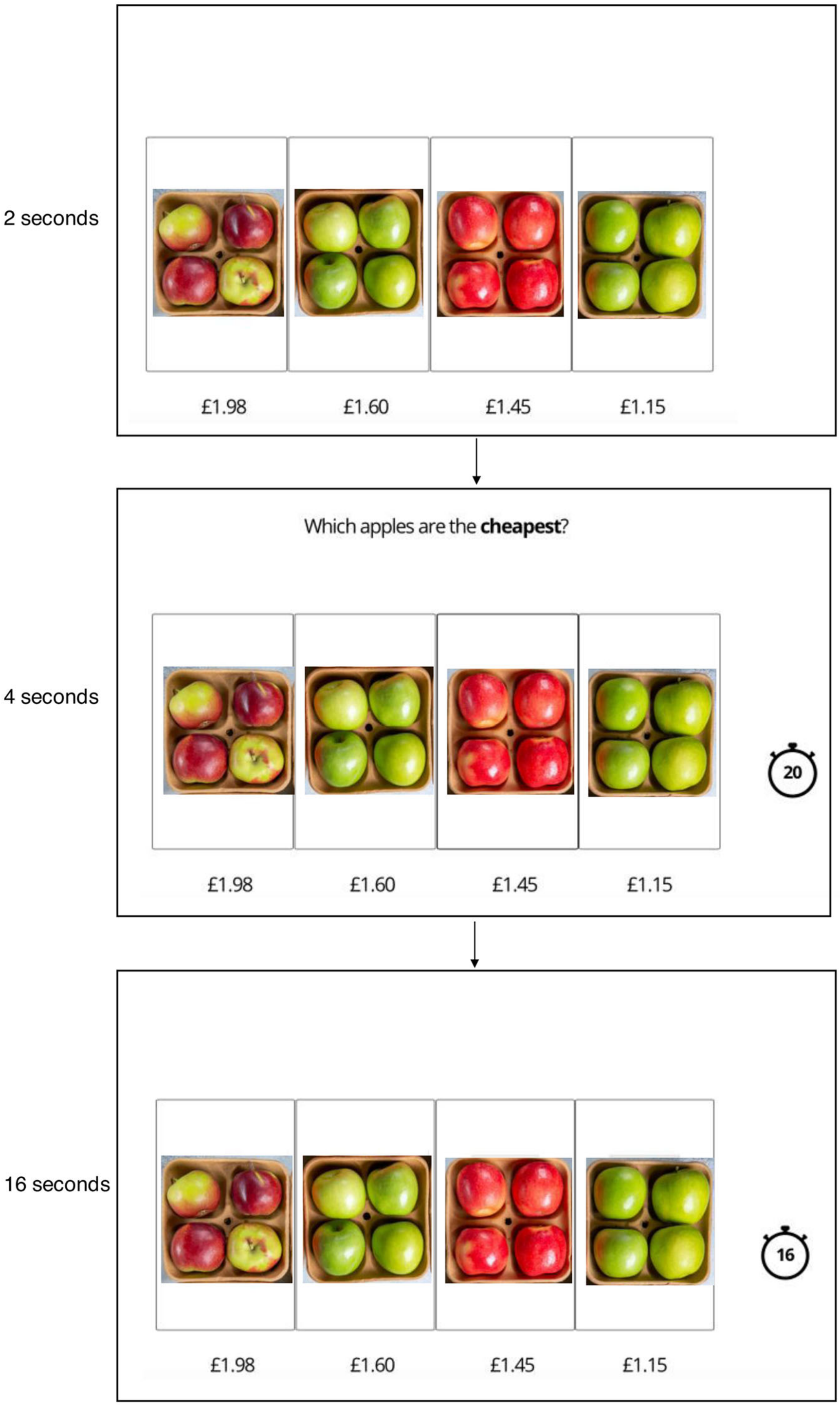

FIGURE 4 Screenshots of question 1 in the TPT showing the timing of each screen (varies per question).
Note: In this question, participants select their answer by clicking the apples.
[Colour figure can be viewed at wileyonlinelibrary.com]

Aphasia-friendly General Health Numeracy Test (aGHNT-6)

The six-item GHNT (GHNT-6; Osborn et al., 2013) is a valid and reliable objective measurement of health numeracy, a subset of functional numeracy. It assesses various numerical skills in the context of health decision-making and was normed on 205 neurotypical adults (M = 42% correct, SD = 30%; KR-20 = 0.77). In this study, it served as a reference measure of objective functional numeracy to validate the FNA23.

The aGHNT-6 was created for the present study based on an aphasia-friendly version of the GHNT-21 (developed for a student project; Esteban Serna, 2020). Adaptations included reduced language demands, emboldened keywords, visuals and audio (Figure 6). Question 1 was

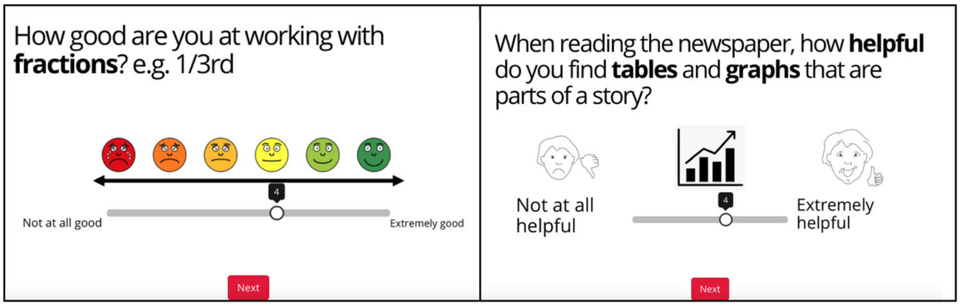

FIGURE 5 Screenshots from the post-stroke SNS measuring current subjective functional numeracy.
Note: Participants register a response by moving the slider.
[Colour figure can be viewed at wileyonlinelibrary.com]

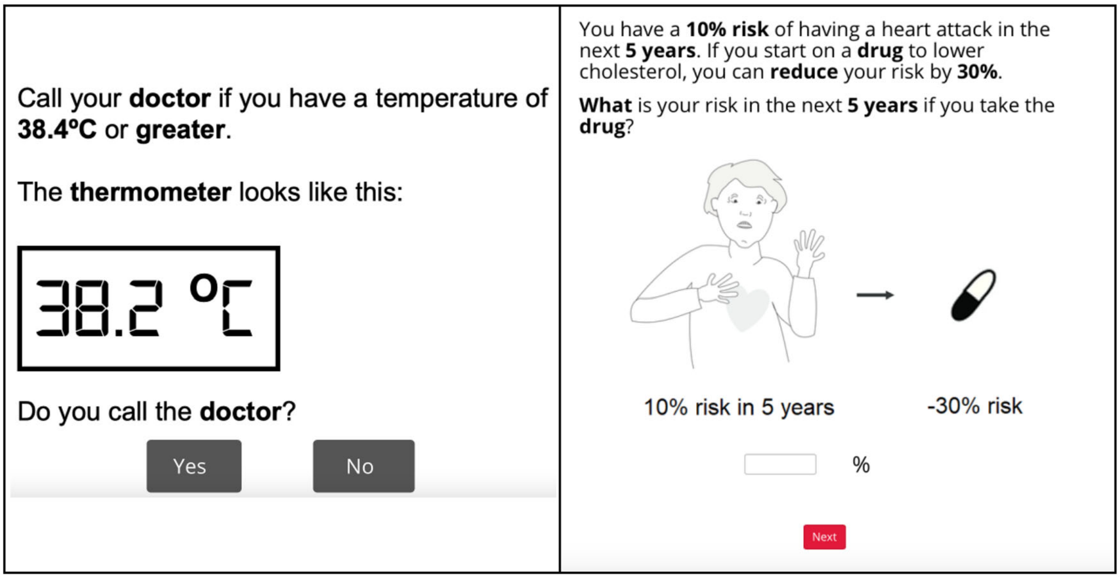

FIGURE 6 Screenshots from aGHNT-6 questions 1 and 5. [Colour figure can be viewed at wileyonlinelibrary.com]

adapted for UK participants by changing the temperatures from Fahrenheit to Celsius, maintaining the required calculation with the same decimals. Gorilla recorded participants' time to complete the task and automatically scored answers as correct or incorrect, resulting in a total out of six, with higher scores indicating stronger functional numeracy.

Card Sorting Task

In recognition of the cognitive flexibility required by the design of the FNA23 (i.e., frequent shifts in topic and response modality), a Card Sorting Task was included in the present study to examine whether PWA's performance on the FNA23 was confounded by their cognitive flexibility. The Card Sorting Task was developed using an existing experiment frame on Gorilla which was inspired by the Wisconsin Card Sorting Test (Berg, 1948).

In each trial, participants were shown four cards across the top of the screen which contained geometric forms that varied in shape, colour and number (Figure 7). Participants were instructed to sort a fifth card by clicking on one of the four cards at the top which matched in shape, colour or number, however they were not explicitly told the sorting rule which changed during the task. Feedback was displayed immediately after each trial indicating whether participants had made a correct match, allowing them to figure out the sorting rule by trial and error. Before starting, participants were shown examples of how to match a card by shape, colour or number. There were 45 trials in total and the rule changed every four to six cards. Gorilla automatically recorded total percentage correct scores for each participant; with higher scores indicating

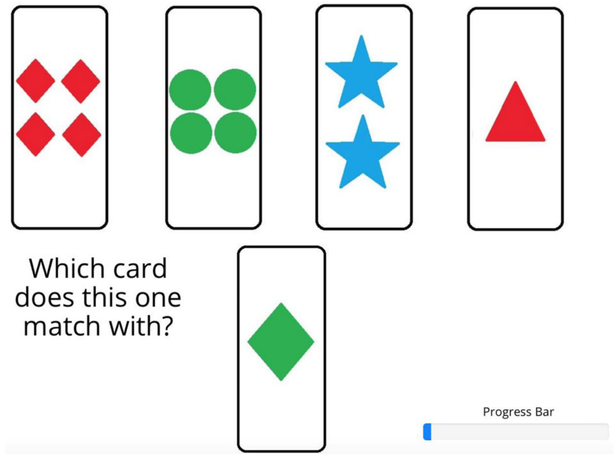

FIGURE 7 Screenshot of a card sorting task trial. [Colour figure can be viewed at wileyonlinelibrary.com]

better cognitive flexibility and thus stronger ability to successfully shift between different task requirements.

Design

The study used within- and between-subjects designs. The variables were total scores and time to complete the FNA23, Time Pressure Task, SNS (pre- and post-stroke for PWA) and aGHNT-6. Aphasia quotient (AQ), age, highest level of education and time post-stroke were additional variables for PWA. Tasks were presented in a fixed order as part of an experimental battery, with no counterbalancing or randomization, reflecting protocols of widely used assessment batteries for PWA, such as the WAB-R.

Procedure

PWA completed three sessions lasting approximately 30–45 min each. Figure 8 contains details of each session's contents. PWA were encouraged to take a break at each midpoint and complete the sessions on different days to limit fatigue, however, most completed the two Gorilla sessions on the same day. Controls completed one Gorilla session lasting approximately 30 min (Figure 8). Participants used hyperlinks to access the Gorilla sessions. They were instructed not to use calculators, and there was an option to test their audio to ensure it recorded correctly for spoken questions. All instructions were presented as text and audio. All tasks were self-paced except the Time Pressure Task. Before starting the Time Pressure Task, participants were informed that they should answer as quickly as possible as questions would be timed, and they would see a clock counting down the remaining seconds. A second instruction screen informed participants that they would see pictures first and then hear and see the question. This screen had a 10-s countdown to familiarize participants with the clock and induce feelings of time pressure.

All data was analysed using IBM SPSS Statistics 28. The SPSS data and syntax that support the findings of this study are openly available on the Open Science Framework (https://osf.io/ckeb8).

RESULTS

Descriptive statistics

PWA ($n = 20$) and controls' ($n = 102$) performance on each of the key measures is summarized in Table 2.

Box plots highlighted multiple potential outliers across the data set. Variables with outliers outside of three standard deviations from the mean included AQ (1 outlier) and time to complete variables for the Time Pressure Task (two in PWA group, one in control group), pre-stroke SNS (two PWA), aGHNT-6 (one PWA, two controls) and FNA23 (one control). Exclusion of the time to complete outliers did not change any conclusions from our analyses; however, exclusion of the AQ outlier did (z-score = –3.63), so analyses involving AQ are reported with this outlier excluded. All other analyses are reported with outliers included.

Shapiro–Wilk tests indicated that most variables were not normally distributed ($p < 0.05$), except PWA's FNA23 scores, post-stroke SNS ratings, Time Pressure Task and Card Sorting Task scores. AQ became normally distributed after exclusion of the outlier. Non-parametric tests were used for analyses involving these non-normal variables.

To reduce the effects of multiple comparisons, a Benjamini-Hochberg (1995) correction was applied to all analyses, using a false discovery rate of 0.1. All significant analyses reported below remained significant after the correction so original (unadjusted) p values are reported for all analyses.

Internal consistency reliability and concurrent validity of the FNA23

Considering the dichotomous scoring of the FNA23, the Kuder–Richardson (*KR-20*) coefficient was calculated to examine the assessment's internal consistency reliability based on PWA's scores. *KR-20* values generally range from 0.0 to 1.0, with values > 0.7 considered acceptable (Salkind, 2010). The internal consistency reliability of the FNA23 was acceptable (*KR-20* = 0.813).

To determine the concurrent validity of the FNA23, correlations were conducted to examine potential

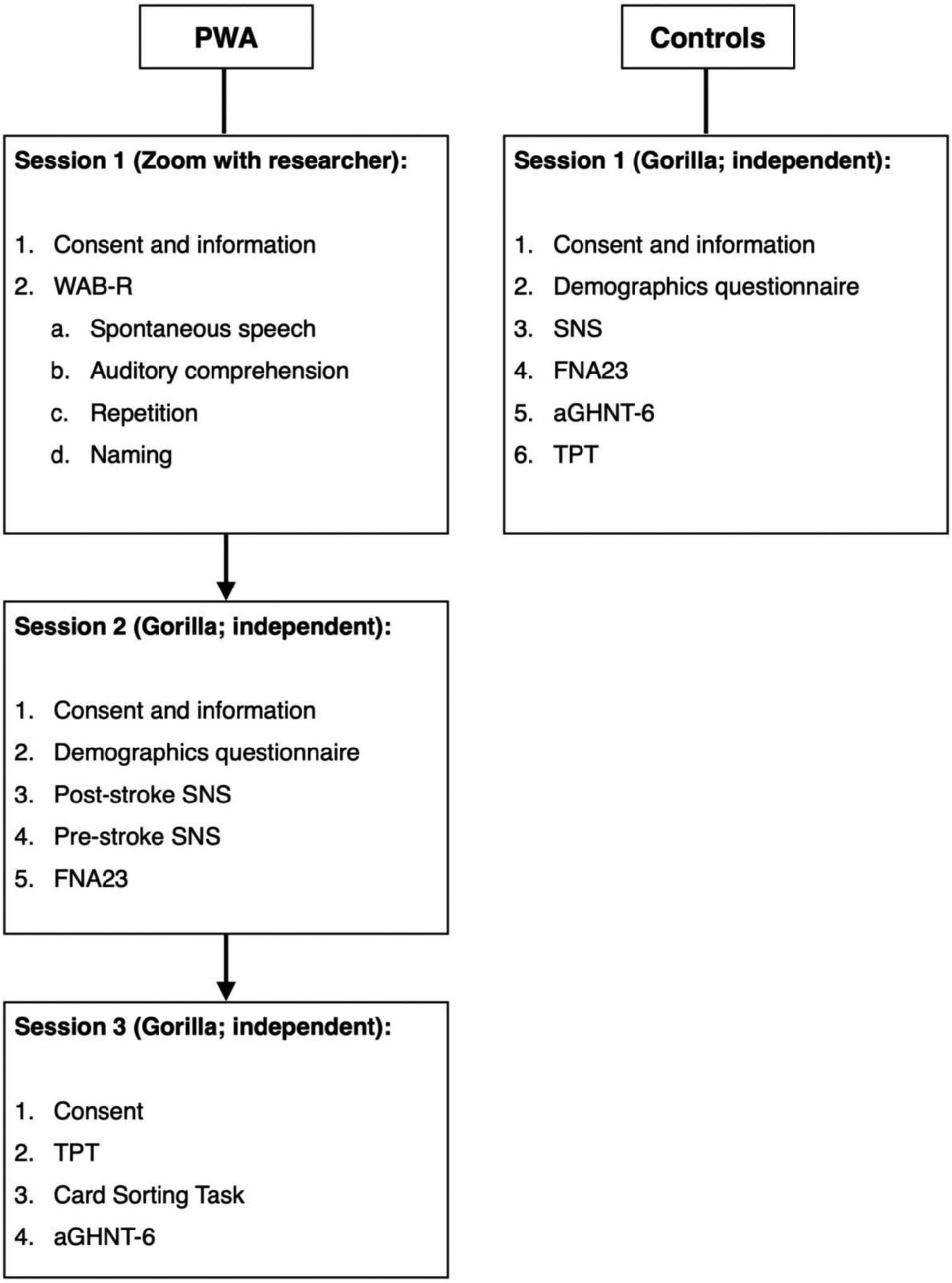

FIGURE 8 Sessions completed by PWA and controls.

relationships between PWA's scores on the functional numeracy measures (Table 3). Pearson correlation coefficients were calculated for correlations involving normally distributed variables (FNA23, Time Pressure Task, post-stroke SNS), and Spearman's Rho was used for correlations involving the non-normally distributed aGHNT-6 scores. Post-hoc power calculations, using G*Power (Faul et al., 2007), revealed that only correlations with a coefficient > 0.59 were sufficiently powered.

Influence of time pressure on PWA's functional numeracy performance

To explore the impact of time pressure, a paired samples *t*-test was conducted to compare PWA's scores on the six matched Time Pressure Task and FNA23 items. The results indicated that PWA's scores on these matched questions were not significantly different ($t(19) = 1.473$, $p = 0.157$, power $= 0.288$).

TABLE 2 Descriptive statistics

	Variable	Mean (SD)	Range	
			Minimum	Maximum
FNA23	*Total score (out of 23)*			
	PWA	15.9 (4.41)	4	22
	Controls	20.22 (2.11)	11	23
	Time to complete (min)			
	PWA	23.97 (17.44)	9.46	69.28
	Controls	8.68 (3.3)	3.71	27.2
	Six questions matched to TPT (out of 6)			
	PWA	4.45 (1.23)	2	6
	Controls	5.62 (0.58)	4	6
TPT	*Total score (out of 7)*			
	PWA	4.5 (1.96)	0	7
	Controls	6.42 (0.72)	4	7
Time to complete (min)	*Time to complete (min)*			
	PWA	5.4 (5.62)	1.71	21.25
	Controls	1.4 (0.37)	0.86	3.21
Six questions matched to FNA (out of 6)	*Six questions matched to FNA (out of 6)*			
	PWA	3.9 (1.77)	0	6
	Controls	5.44 (0.7)	3	6
SNS	*Self-rating (between 1 and 6)*			
	PWA pre-stroke	4.99 (0.87)	3.38	6
	PWA post-stroke	3.95 (0.96)	2	6
	Controls	4.63 (0.92)	1	6
	Time to complete (min)			
	PWA pre-stroke	2.45 (1.87)	0.55	7.55
	PWA post-stroke	4.11 (2.64)	1.41	13.08
	Controls	2 (0.6)	0.8	3.74
aGHNT-6	*Total score (out of 6)*			
	PWA	2.3 (1.34)	1	5
	Controls	4.12 (1.34)	1	6
	Time to complete (min)			
	PWA	7.83 (7.03)	3.01	33.19
	Controls	3.97 (1.87)	1.35	11.67
Card sorting task	*Per cent correct*			
	PWA	42.92 (9.16)	20	62

TABLE 3 Correlations between PWA's scores on the functional numeracy measures

Variable	FNA23	TPT	Post-stroke SNS	aGHNT-6
FNA23[a]	–			
TPT[a]	0.54*	–		
Post-stroke SNS[a]	0.34	0.39	–	
aGHNT-6[b]	0.66*	0.59*	0.51*	–

Notes: [a] Pearson correlations.
[b]Spearman's Rho correlations.
**Significant with the Benjamini–Hochberg correction.

Impact of aphasia on functional numeracy

PWA versus controls

To explore the impact of aphasia on functional numeracy skills, PWA and controls were compared based on their performance on the functional numeracy measures, the FNA23, Time Pressure Task, aGHNT-6 and SNS (post-stroke for PWA), and the time they took to complete each task. Since controls' and PWA's ages were significantly different, analyses were performed both with the entire

TABLE 4 Mann–Whitney *U*-tests comparing PWA and controls' scores on the functional numeracy measures and time taken to complete each task

	Median							
	PWA	Controls	*U*	*z*	*N*	*p*	*r* (effect size)	Power
Total scores								
FNA23	16.5	21	337.5	−4.79	122	< 0.001	0.43 (medium)	> 0.99
TPT	4	7	399.5	−4.66	122	< 0.001	0.42 (medium)	> 0.99
SNS (post-stroke)	3.75	5	594.5	−2.98	122	0.003	0.27 (small)	> 0.99
aGHNT-6	2	4	368.5	−4.61	122	< 0.001	0.42 (medium)	> 0.99
Time to complete (min)								
FNA23	16.65	7.73	147	−6.04	122	< 0.001	0.55 (large)	>0.99
TPT	3.27	1.31	73	−6.55	122	< 0.001	0.59 (large)	0.98
SNS (post-stroke)	3.21	2	343	−4.68	122	< 0.001	0.42 (medium)	> 0.99
aGHNT-6	4.38	3.62	553	−3.23	122	0.001	0.29 (small)	0.87

Note: All comparisons remained significant with the Benjamini–Hochberg correction.

sample and with participants younger than 30 excluded. Excluding the younger participants did not change any conclusions; therefore, results reported below include all participants. Mann–Whitney *U*-tests were used since controls' variables were all not normally distributed. Results indicated that, compared with controls, PWA scored significantly lower on all measures and took significantly longer to complete the tasks (Table 4).

TABLE 5 Correlations between PWA's performance on the functional numeracy measures and AQ

Variable	AQ
FNA23[a]	0.48*
TPT[a]	0.61*
Post-stroke SNS[a]	0.23
aGHNT-6[b]	0.51*

Note: [a] Pearson correlations. [b] Spearman's Rho correlations. *

Subjective impact of aphasia on functional numeracy

PWA's ratings on the pre- and post-stroke SNS were compared with explore the subjective impact of aphasia on their functional numeracy abilities and preferences. A Wilcoxon signed-rank test was used because the data for the pre-stroke SNS were negatively skewed. Results indicated that PWA's subjective numeracy ratings were significantly lower on the post-stroke SNS (median = 3.75) compared with the pre-stroke SNS (median = 5.25, $W = 3$, $z = -3.71$, $n = 20$, $p < 0.001$, power = 0.97). The effect size was large ($r = 0.83$). The comparison remained significant with the Benjamini–Hochberg correction.

Relationships between functional numeracy and aphasia severity

To determine whether PWA's functional numeracy skills were related to their aphasia severity, correlations were conducted to examine the relationships between PWA's AQ and performance on the functional numeracy measures (Table 5). There were significant positive correlations between aphasia severity and scores on the FNA23, Time Pressure Task and aGHNT-6, with less severe aphasia (higher AQ) relating to better performance on the tasks. The correlation between AQ and post-stroke SNS ratings was not significant. Again, only correlations > 0.59 were sufficiently powered.

Relationship between functional numeracy and cognitive flexibility

A Pearson Correlation Coefficient was calculated to determine whether PWA's FNA23 scores were related to their cognitive flexibility, as measured by the Card Sorting Task. Results indicated a non-significant (and underpowered) correlation between PWA's FNA23 scores and cognitive flexibility (Pearson's $r = 0.354$, d.f. = 17, $p = 0.137$; 12.5% shared variance).

Predictors of PWA's functional numeracy

A multiple linear regression was conducted to investigate whether demographic factors were significant predictors of PWA's functional numeracy and potentially contributing to a relationship between aphasia severity and functional numeracy. Relevant assumptions were checked, and a regression was run with AQ, age, education (included as dummy variables) and time post-stroke as the predictors (independent variables) and FNA23 scores as the outcome (dependent variable). The regression model was not significant ($F(8,8) = 1.66$, $p = 0.246$, $R^2 = 0.62$).

DISCUSSION

This study aimed to design a wide-ranging, aphasia-friendly functional numeracy assessment, the FNA23, and evaluate its validity and reliability in a pilot group of 20 individuals with aphasia. Additional aims included investigating whether functional numeracy was impacted by time pressure and aphasia severity and exploring predictors of PWA's functional numeracy.

Reliability and validity of the FNA23

The FNA23 demonstrated acceptable internal consistency reliability, perfect interrater reliability for spoken items, and a significantly strong correlation with the objective functional numeracy measure used to examine concurrent validity (aGHNT-6). Contrasting with their positively skewed aGHNT-6 scores (indicating a floor effect), PWA's FNA23 scores were normally distributed. These results suggest that the FNA23's language load and difficulty level are suitable for assessing PWA with varying abilities and that the FNA23 can be considered a valid and reliable objective measure of PWA's functional numeracy.

The FNA23 and SNS were not significantly correlated, indicating distinctions between PWA's objective and subjective functional numeracy. This supports Liberali et al.'s (2012) argument that objective and subjective numeracy measures do not measure identical constructs. Perhaps some PWA had sufficient ability to complete objective numeracy tasks, thus scored well on the FNA23, but perceived themselves as less proficient because of the time and effort required. It has also been reported that PWA may perceive their performance as worse than it is (Langland-Hassan et al., 2017), which may explain these findings. Carvalho (2015) suggested that future research with the SNS would benefit from investigating people in the lowest numeracy bracket, who are likely more aware of their difficulties. The SNS may, therefore, not be as sensitive as the FNA23 in detecting PWA's functional numeracy difficulties as they mostly fell into these lower numeracy brackets. Additionally, although the SNS has been reported to correlate with objective functional numeracy measures (e.g., Osborn et al., 2013; Zikmund-Fisher et al., 2007), this study's contradictory findings and low power may reflect differences in the numeracy measures used.

Influence of time pressure on PWA's functional numeracy

Contrary to predictions, PWA did not score significantly differently on the matched FNA23 and Time Pressure Task questions, suggesting the addition of time pressure did not have a significant effect. This was surprising considering PWA's commonly reported slowed processing speeds and working memory impairments (e.g., Faroqi-Shah & Gehman, 2021; Laures-Gore et al., 2011), which were hypothesized to disadvantage PWA in time pressured situations that likely place additional demands on these processes (Caviola et al., 2017; Echt et al., 1998).

Orfus's (2008) study of the effects of time pressure on undergraduates' mathematics performance also did not find a significant effect of time pressure which they attributed to their task creating insufficient feelings of time pressure and seriousness. Despite the Time Pressure Task's time limits being based on neurotypicals' pilot data, perhaps they were too long to induce sufficient feelings of pressure to place additional demands on processing speed or working memory. However, the absence of significant findings in this case and elsewhere in this study should be interpreted with caution due to the within-subject analyses being underpowered.

Future research should investigate appropriate time limits which induce pressure but avoid a floor effect. As suggested by Orfus (2008), participants should be asked after the task whether they felt time pressured to ascertain whether time pressure was sufficiently induced. Increasing participants' motivation through including an incentive of importance (e.g., through gamification; Cerrato & Ponticorvo, 2017) may also increase the effects of time pressure (Orfus, 2008). If time pressure is found to impact PWA's performance, adding time pressure to part of the FNA23 may increase its ecological validity.

Impact of aphasia on functional numeracy

As a group, PWA demonstrated significantly poorer functional numeracy than controls and took significantly longer on all functional numeracy measures, evidencing the prevalence of functional numeracy difficulties in PWA. However, as in Villain et al.'s (2015) study, some PWA scored similarly to controls but still took longer to complete the tasks, indicating they would likely be disadvantaged in everyday situations. This highlights the necessity of considering both PWA's scores and timing when assessing their functional numeracy.

Although controls were not matched to PWA's demographics, findings did not change when analyses were rerun with age-matched groups. However, it is a limitation that other factors which may have affected functional numeracy, such as socioeconomic status and education, were not considered (e.g., Smith et al., 2010).

Comparing PWA's pre- and post-stroke SNS ratings suggested their subjective functional numeracy abilities and preferences were negatively affected by their stroke, provided PWA's retrospective ratings were accurate. Although these ratings cannot be verified by objective tests, Cooper-Evans et al. (2008) argued that retrospective ratings were a useful measure of people with acquired brain injuries' subjective perceptions of changes since their injury.

Together, these findings support the hypothesis that aphasia impacts functional numeracy, concurring with Proios et al. (2021) and Robert et al.'s (2021) findings using the NADL and EABN. Importantly, this study's findings were based on assessments designed to be accessible for PWA. This suggests that the language demands of the NADL and EABN were likely sufficiently low not to confound PWA's performance.

Relationship between functional numeracy and aphasia severity

Consistent with Robert et al.'s (2021) finding that more severe aphasia correlated with poorer EABN scores, this study found that more severe aphasia correlated with poorer objective functional numeracy as assessed by the FNA23, Time Pressure Task and aGHNT-6. However, aphasia severity was not correlated with subjective functional numeracy abilities and preferences. These findings partly support the hypothesis that functional numeracy is associated with aphasia severity and again indicate a distinction between PWA's objective and subjective functional numeracy.

This study did not find that AQ, age, education or time post-stroke significantly predicted PWA's scores on the FNA23. The non-significant contributions of AQ, age and time post-stroke were in line with Gonzalez et al. (2021) findings. However, the non-significant contribution of education to the regression model was surprising given Gonzalez et al.'s findings and Proios et al.'s (2021) discussion about the possible protective role of education on both formal and functional mathematical abilities. However, this null finding perhaps resulted from this study considering participants' highest level of education rather than years in education, which may have provided a more sensitive measure (Ritchie & Tucker-Drob, 2018). The low power of this study's regression analysis may also have contributed to a potential false negative result regarding education.

Given reports of cognitive flexibility impairments in PWA (e.g., Spitzer et al., 2020), it was recognized that the frequent shifts in topic and response modality required by the FNA23's design might have confounded PWA's performance. However, the non-significant correlation between PWA's scores on the FNA23 and Card Sorting Task suggested that cognitive flexibility impairments were unlikely to have affected PWA's performance on the FNA23.

Future research should consider additional cognitive and language factors (e.g., syntactic comprehension), which might predict PWA's functional numeracy (Boyle et al., 2013; Nakai & Okanoya, 2018) and contribute to the relationship between AQ and functional numeracy. This might highlight specific skills that could have wide-ranging effects on functional numeracy if prioritized in interventions.

This study did not have the power to conduct in-depth analyses of PWA's performance across different skills and response modalities in the FNA23. However, informal observations indicated that PWA generally performed more poorly on transcoding, number comparison and calculation questions and better on number identification and counting. Typed responses appeared most difficult, and multiple-choice picture responses the easiest. Future research should examine PWA's FNA23 performance more closely to determine whether these differences in performance across different skills and response modalities are related to aphasia severity or perhaps impacted differently by certain aphasia subtypes, as might be expected by Gonzalez et al. (2021). Such a study requires more systematic recruitment of participants with different aphasia severities and subtypes than was possible in the current study.

Limitations and future directions

Although this study provided a successful pilot validation of the FNA23 and replicated previous findings regarding functional numeracy difficulties in PWA, there are potential limitations regarding recruitment, the FNA23's design and its unsupervised delivery.

Conducting the study remotely and entirely online was advantageous as participants could be recruited from across the country without needing to travel, saving time and expenses. However, only participants who were able to access the internet via smartphones, tablets or computers were included in the study. This inclusion criterion was necessary to reduce potential effects of poor technological skills confounding participants' performance but might have caused a selection bias excluding participants with a lower socioeconomic status and thus potentially poorer numeracy skills (e.g., Smith et al., 2010). Considering controls and PWA were both affected by this inclusion criterion, this bias is unlikely to have significantly influenced between-subjects findings; however, it does limit the findings' generalizability to the wider population. More broadly, this bias may have implications for other aphasia research now conducted online since COVID-19 (Kong, 2021). This may be addressed by providing technology and technological support to PWA who need it, rather than excluding them, and by recording participants' socioeconomic status to objectively assess whether the bias exists.

Although the FNA23 aimed to accommodate poor reading ability by including audio with typed questions, some stimuli contained crucial text that was not accompanied by audio and thus might have confounded performance. This limitation was exacerbated by not having an assessor present to prompt participants when necessary. Future studies should consider participants' reading abilities and potentially include audio descriptions of stimuli and emboldened keywords to increase accessibility (Rose et al., 2011).

The FNA23's computerized format was advantageous in enabling identical delivery and automatic scoring (except for spoken questions). However, further research should consider supervising its delivery (face to face or over video call) to limit potential interference from technological difficulties and enable observational data on PWA's experiences to be collected, which may aid future development of the assessment. Supervising delivery is also more beneficial in clinical settings so clinicians can prompt PWA when necessary, observe self-corrections and note any strategies used to complete the tasks, which will assist speech and language therapists to select appropriate interventions. Further research should also consider whether there is a difference in PWA's performance if they use a smartphone, tablet or computer to complete the FNA23.

Development of the FNA23 is still in its infancy and given the relatively small sample of PWA, many within-subjects comparisons in this study were underpowered. Future studies are needed with larger samples of PWA representing various aphasia severities and subtypes. Interrater reliability for spoken questions should be re-evaluated for people with more severe aphasia as more specific scoring criteria may need to be developed. Test–retest reliability needs to be investigated, and cut-off timings and scores should be calculated based on controls' performance, considering factors (e.g., education) that may affect performance (e.g., following methods used by Semenza et al., 2014). These cut-offs can be used to investigate the prevalence of PWA's functional numeracy difficulties more systematically.

Implications and conclusions

This pilot validation study of the FNA23 fills a gap in the literature by reporting findings from the first valid, reliable and wide-ranging functional numeracy assessment suitable for PWA. The FNA23 is sensitive enough to detect a range of functional numeracy abilities, and with further development, it will be a useful tool in both research and clinical practice to evaluate PWA's functional numeracy strengths and difficulties to inform intervention design and monitor progress towards achieving intervention objectives. This study was an important first step in the FNA23's development and adds to the growing body of evidence highlighting the prevalence of functional numeracy difficulties in PWA, the relationship between functional numeracy and aphasia severity, and most importantly, the need for PWA to be evaluated for functional numeracy difficulties as part of their neurorehabilitation journeys.

ACKNOWLEDGEMENTS

The authors are very grateful to all the participants who took part in this project.

PATIENT CONSENT STATEMENT

Informed consent was obtained from all participants before beginning the study.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available on the Open Science Framework (https://osf.io/ckeb8).

ORCID

Kerri Ichikowitz https://orcid.org/0000-0002-9015-6708

REFERENCES

Benjamini, Y. & Hochberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57(1), 289–300.

Benn, Y., Jayes, M., Casassus, M., Williams, M., Jenkinson, C., McGowan, E. & Conroy, P. (2022) A qualitative study into the experience of living with acalculia after stroke and other forms of acquired brain injury. *Neuropsychological Rehabilitation*, 1–25. Advance online publication. https://doi.org/10.1080/09602011.2022.2108065

Berg, E.A. (1948) A simple objective technique for measuring flexibility in thinking. *The Journal of General Psychology*, 39(1), 15–22.

Boyle, P.A., Yu, L., Wilson, R.S., Segawa, E., Buchman, A.S. & Bennett, D.A. (2013) Cognitive decline impairs financial and health literacy among community-based older persons without dementia. *Psychology and Aging*, 28(3), 614.

Brennan, A., Worrall, L. & McKenna, K. (2005) The relationship between specific features of aphasia-friendly written material and comprehension of written material for people with aphasia: an exploratory study. *Aphasiology*, 19(8), 693–711.

Burgess, P.W., Alderman, N., Forbes, C., Costello, A., Laure, M.C., Dawson, D.R., Anderson, N.D., Gilbert, S.J., Dumontheil, I. & Channon, S. (2006) The case for the development and use of "ecologically valid" measures of executive function in experimental and clinical neuropsychology. *Journal of the International Neuropsychological Society*, 12(2), 194–209.

Burgio, F., Danesin, L., Benavides-Varela, S., Meneghello, F., Butterworth, B., Arcara, G. & Semenza, C. (2022) Numerical activities of daily living: a short version. *Neurological Sciences*, 43(2), 967–978.

Carvalho, A. (2015) *A self-report measure of number sense. [Unpublished masters thesis]*. Fresno: California State University.

Caviola, S., Carey, E., Mammarella, I.C. & Szucs, D. (2017) Stress, time pressure, strategy selection and math anxiety in mathematics: a review of the literature. *Frontiers in Psychology*, 8, 1488.

Cerrato, A. & Ponticorvo, M. (2017) Enhancing neuropsychological testing with gamification and tangible interfaces: the baking tray task. *Biomedical Applications Based on Natural and Artificial Computing*, 147–156.

Cooper-Evans, S., Alderman, N., Knight, C. & Oddy, M. (2008) Self-esteem as a predictor of psychological distress after severe acquired brain injury: an exploratory study. *Neuropsychological Rehabilitation*, 18(5–6), 607–626.

Data Protection Act. (2018) Viewed 15 November, 2020, https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted

Dekhtyar, M., Braun, E.J., Billot, A., Foo, L. & Kiran, S. (2020) Videoconference administration of the Western aphasia battery–revised: feasibility and validity. *American Journal of Speech–Language Pathology*, 29(2), 673–687.

Delazer, M., Girelli, L., Semenza, C. & Denes, G. (1999) Numerical skills and aphasia. *Journal of the International Neuropsychological Society*, 5(3), 213–221.

Echt, K.V., Morrell, R.W. & Park, D.C. (1998) Effects of age and training formats on basic computer skill acquisition in older adults. *Educational Gerontology: An International Quarterly*, 24(1), 3–25.

Esteban Serna, C. (2020) *What counts to live well with acquired speech impairment: Health numeracy as a mediator between working memory and quality of life in people with aphasia [Unpublished undergraduate dissertation]*. University College London.

Excellence Gateway. (n.d.) *Numeracy* . Viewed 15 November 2020, https://www.excellencegateway.org.uk/interactive-resources/numeracy

Fagerlin, A., Zikmund-Fisher, B.J., Ubel, P.A., Jankovic, A., Derry, H.A. & Smith, D.M. (2007) Measuring numeracy without a math test: development of the subjective numeracy scale. *Medical Decision Making*, 27(5), 672–680.

Faroqi-Shah, Y. & Gehman, M. (2021) The role of processing speed and cognitive control on word retrieval in aging and aphasia. *Journal of Speech, Language, and Hearing Research*, 64(3), 949–964.

Faul, F., Erdfelder, E., Lang, A.G. & Buchner, A. (2007) G* Power 3: a flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191.

Frattali, C., Thompson, C.M., Holland, A.L., Wohl, C.B. & Ferketic, M.M. (1995) *Functional assessment of communication skills for adults (ASHA-FACS)*. Rockville, MD: American Speech–Language–Hearing Association.

Fromm, D., Forbes, M., Holland, A., Dalton, S.G., Richardson, J. & MacWhinney, B. (2017) Discourse characteristics in aphasia beyond the Western Aphasia Battery cutoff. *American Journal of Speech–Language Pathology*, 26(3), 762–768.

Ginsburg, L., Manly, M. & Schmitt, M.J. (2006) The Components of Numeracy. *NCSALL Occasional Papers*. Cambridge, MA: National Center for the Study of Adult Learning and Literacy.

Gonzalez, R., Rojas, M., Rosselli, M. & Ardila, A. (2021) Acalculia in aphasia. *Archives of Clinical Neuropsychology*, 36(4), 455–464.

Holland, W.L. & Fromm, D. (2018) *CADL-3: Communication activities of daily living*, 3rd edition, Austin, Tex: Pro-Ed.

Howe, T.J., Worrall, L.E. & Hickson, L.M. (2008) Observing people with aphasia: environmental factors that influence their community participation. *Aphasiology*, 22(6), 618–643.

Kertesz, A. (2007) *Western aphasia battery–R*. New York, NY: Grune & Stratton.

Kong, A.P.H. (2021) The impact of COVID-19 on speakers with aphasia: what is currently known and missing? *Journal of Speech, Language, and Hearing Research*, 64(1), 176–180.

Langland-Hassan, P., Gauker, C., Richardson, M.J., Dietz, A. & Faries, F.R. (2017) Metacognitive deficits in categorization tasks in a population with impaired inner speech. *Acta Psychologica*, 181, 62–74.

Laures-Gore, J., Marshall, R.S. & Verner, E. (2011) Performance of individuals with left hemisphere stroke and aphasia and individuals with right brain damage on forward and backward digit span tasks. *Aphasiology*, 25(1), 43–56.

Liberali, J.M., Reyna, V.F., Furlan, S., Stein, L.M. & Pardo, S.T. (2012) Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making*, 25(4), 361–381.

Lipkus, I.M. & Peters, E. (2009) Understanding the role of numeracy in health: proposed theoretical framework and practical insights. *Health Education & Behavior*, 36(6), 1065–1081.

Nakai, T. & Okanoya, K. (2018) Neural evidence of cross-domain structural interaction between language and arithmetic. *Scientific Reports*, 8(1), 1–9.

Orfus, S. (2008) The effect test anxiety and time pressure on performance. *The Huron University College Journal of Learning and Motivation*, 46(1).

Osborn, C.Y., Wallston, K.A., Shpigel, A., Cavanaugh, K., Kripalani, S. & Rothman, R.L. (2013) Development and validation of the general health numeracy test (GHNT). *Patient Education and Counseling*, 91(3), 350–356.

Pearson Education. (2015) *Assessing and maintaining test quality*. Viewed 30 August, 2021, https://www.pearsonvue.co.uk/Documents/Security/pearson_vue_test_quality_US.aspx

Proios, H., Tsakpounidou, K., Karapanayiotides, T., Priftis, K. & Semenza, C. (2021) Aphasia and math: deficits with basic number comprehension and in numerical activities of daily living. *Journal of the International Neuropsychological Society*, 27(9), 939–951.

Ritchie, S.J. & Tucker-Drob, E.M. (2018) How much does education improve intelligence? A meta-analysis. *Psychological science*, 29(8), 1358–1369.

Robert, H., Villain, M., Prevost-Tarabon, C., Cocquelet-Bunting, M., Glize, B., Pradat-Diehl, P. & Bayen, E. (2021) Ecological assessment of numerical skills in adults with left stroke. *Annals of Physical and Rehabilitation Medicine*, 64(1), 101383.

Rose, T.A., Worrall, L.E., Hickson, L.M. & Hoffmann, T.C. (2011) Aphasia friendly written health information: content and design characteristics. *International Journal of Speech–Language Pathology*, 13(4), 335–347.

Salkind, N.J. (Ed.) (2010) *Encyclopedia of research design*, vol. 1. Los Angeles, Calif.:SAGE.

Semenza, C. (2008) Number processing. *Handbook of the neuroscience of language*, London: Academic Press, 219–227.

Semenza, C., Meneghello, F., Arcara, G., Burgio, F., Gnoato, F., Facchini, S., Benavides-Varela, S., Clementi, M. & Butterworth, B. (2014) A new clinical tool for assessing numerical abilities in neurological diseases: numerical activities of daily living. *Frontiers in Aging Neuroscience*, 6, 112.

Smith, S.G., Wolf, M.S. & Wagner, C.V. (2010) Socioeconomic status, statistical confidence, and patient–provider communication: an analysis of the Health Information National Trends Survey (HINTS 2007). *Journal of Health Communication*, 15(sup 3), 169–185.

Spitzer, L., Binkofski, F., Willmes, K. & Bruehl, S. (2020) Executive functions in aphasia: a novel aphasia screening for cognitive flexibility in everyday communication. *Neuropsychological Rehabilitation*, 30(9), 1701–1719.

Stroke Association. (2017) *State of the nation stroke statistics*. Viewed 7 May, 2022, https://www.mynewsdesk.com/uk/stroke-association/documents/state-of-the-nation-2017-68765

Swinburn, K., Porter, G. & Howard, D. (2004) Comprehensive aphasia test. Hove: Psychology Press.

Villain, M., Tarabon-Prevost, C., Bayen, E., Robert, H., Bernard, B., Hurteaux, E. & Pradat-Diehl, P. (2015) Ecological Assessment Battery for Numbers (EABN) for brain-damaged patients: standardization and validity study. *Annals of Physical and Rehabilitation Medicine*, 58(5), 283–288.

Wallace, S.J., Worrall, L., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., Hin Kong, A.P., Simmons-Mackie, N., Scarinci, N. & Gauvreau, C.A. (2017) 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.

Zikmund-Fisher, B.J., Smith, D.M., Ubel, P.A. & Fagerlin, A. (2007) Validation of the Subjective Numeracy Scale: effects of low numeracy on comprehension of risk communications and utility elicitations. *Medical Decision Making*, 27(5), 663–671.

How to cite this article: Ichikowitz, K., Bruce, C., Meitanis, V., Cheung, K., Kim, Y., Talbourdet, E. et al. (2023) Which blueberries are better value? The development and validation of the functional numeracy assessment for adults with aphasia. *International Journal of Language & Communication Disorders*, 1–22. https://doi.org/10.1111/1460-6984.12867

APPENDIX A

A1, B1

International Journal of Language & Communication Disorders

TABLE A1 Questions in the FNA23

	Question	Topic	Mode of response	Skill
1.	What time is it?	Digital clock	Spoken	Number identification
2.	How much battery is left on the clock?		Spoken	Number identification
3.	How many draws were there this week?	Sports league tables	MCQ typed	Counting
4.	How many teams did not score in their matches?		MCQ typed	Counting
5.	How many points does the third team have?		Spoken	Number identification
6.	How many wins does Arsenal have?		Spoken	Number identification
7.	Can you tell me Garry's phone number?	Telephone	Spoken	Number identification
8.	Whose number ends with '176'?		MCQ typed	Number identification
9.	Here is Jane's phone number. Can you type it below? Click play to listen to the number. Then click in the text box so you can type it while you listen.		Typed	Transcoding
10.	You have 720 g of flour. How many scones can you bake?	Cooking and baking	Typed	Computation
11.	Which orange juice is the cheapest one?	Shopping	MCQ picture	Number comparison
12.	You have £60. You want 4 bottles of perfume. Which 4 can you get?		MCQ picture	Computation
13.	Which blueberries are better value?		MCQ picture	Computation
14.	Here is a promotion. You buy three items and get the cheapest one for free. Which one of these items do you get for free?		MCQ picture	Number comparison
15.	You are buying these two items. How much will you pay?		Typed	Computation
16.	Your shopping costs £4.20. Which box has enough money?		MCQ picture	Computation
17.	You bought a book and it cost £12.49. You paid with a £20 note. How much will you receive as change?		Typed	Computation
18.	You catch the train at Finsbury Park. You travel seven stations towards Brixton. Where do you get off?	Travel	MCQ typed	Counting
19.	You are at Brixton and need to go to Warren Street. How many stations do you have to travel?		Spoken	Counting
20.	Can you identify the correct Wi-Fi password?	Password	MCQ typed	Number identification
21.	It's 3:25 pm now, the train to Basingstoke leaves at 16:12, how many minutes do you have before the departure?	Travel	Typed	Computation
22.	You are making risotto. You have five guests. How much risotto rice do you need?	Cooking and baking	Typed	Computation
23.	Which city had the highest recorded temperature in February?	Graph	MCQ typed	Number comparison

Note: MCQ, multiple-choice question.

TABLE B1 Questions in the Time Pressure Task

Question	Corresponding FNA question	Mode of response	Skill	Time limit (s)
1. Which apples are the cheapest?	11.	MCQ picture	Number comparison	20
2. You are buying these two items. How much will you pay?	15.	Typed	Computation	30
3. You bought a jumper. It cost £14.49. You paid with a £20 note. How much will you receive as change?	17.	Typed	Computation	30
4. Please tell me Anne's phone number.	7.	Spoken	Number identification	30
5. Whose number ends with '167'?	8.	MCQ typed	Number identification	20
6. Click play to listen to Tom's phone number. Then click in the textbox so you can type it while you listen.	9.	Typed	Transcoding	30
7. This till is for 10 items or less. Which basket can you take to this till?	–	MCQ picture	Counting	35

Note: MCQ, multiple-choice question.