BMJ Open

Adaptive balancing of effort, accuracy and response speed in anomia treatment for post-stroke aphasia in community-based settings in the USA: a within-subjects randomised controlled trial protocol

Candace M van der Stelt ,[1] Robert Cavanaugh ,[2] William D Hula ,[1,3] Jeffrey Starns,[4] Alyssa Kelly ,[1] Mara L Goodman ,[1] Lauren Terhorst ,[1] Mohammad Hassany ,[5] Peter Brusilovsky ,[5] William S Evans [1]

To cite: van der Stelt CM, Cavanaugh R, Hula WD, *et al*. Adaptive balancing of effort, accuracy and response speed in anomia treatment for post-stroke aphasia in community-based settings in the USA: a within-subjects randomised controlled trial protocol. *BMJ Open* 2026;**16**:e116172. doi:10.1136/bmjopen-2025-116172

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (https://doi.org/10.1136/bmjopen-2025-116172).

Received 31 December 2025
Accepted 21 April 2026

For numbered affiliations see end of article.

Correspondence to
Dr William S Evans;
will.evans@pitt.edu

ABSTRACT

Introduction Anomia is a primary feature of aphasia that negatively impacts quality of life. Although current anomia treatments improve word retrieval, long-term retention and generalisation of trained words to discourse-level communication are rarely measured. Treatment that produces lasting naming gains and generalises to real-world use is one of the top priorities of people living with aphasia. Here, we report the protocol for a randomised clinical trial that investigates individualised anomia treatment through adaptive naming deadlines to achieve 'desirable difficulty' to promote learning retention and generalisation.

Methods and analysis We implement a within-subject sequential, crossover design in which 30 participants with chronic post-stroke aphasia will complete three anomia treatment conditions in randomised order: (1) an adaptive condition where the naming deadline (ie, amount of time the participant is given to name the item) dynamically adjusts between 1.5 and 10 s based on ongoing participant performance and (2) a static Effort-Maximised condition where there is a fixed 10-second naming deadline for all treatment sessions and (3) a static Accuracy-Maximised condition where items are presented immediately in auditory and orthographic form and are repeated by the participant. In each condition, participants are treated on 40 unique non-overlapping words across eight treatment sessions. Before and after each condition, participants complete naming probes and discourse probes. Treatment outcomes from the adaptive treatment will be tested against the two static conditions using linear mixed-effects modelling. Our primary outcome is performance on noun picture naming at 3 months post-treatment. We evaluate production of treated words in discourse probes as a secondary analysis. We predict that our novel, adaptive naming treatment will produce more successful outcomes compared with the static treatment conditions.

Ethics and dissemination The Institutional Review Board of the University of Pittsburgh approved the trial protocol (Study 21120130). Following study completion, results will be disseminated in peer-reviewed journals. If hypothesised results are observed, the adaptive treatment will be a novel, empirically based intervention for long-term retention of anomia treatment gains, positively impacting the lives and recovery of individuals living with aphasia.

Trial registration number NCT05653440.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This randomised clinical trial measures long-term retention and context generalisation of trained words, which are typically not measured in anomia treatment research.
⇒ The novel adaptive treatment applies principles of learning theory to provide precision rehabilitation, which is predicted to improve the durability and generalisation of treatment gains.
⇒ Blinding participants and treating research assistants to treatment condition is not feasible due to the nature of the behavioural interventions; we counteract this limitation by blinding research assistants who will score probe sessions for final analyses to both time point and condition.
⇒ Participant retention may be challenged by the long-term nature of the study (ie, 10-month average duration).

BACKGROUND

Aphasia is a language impairment that most commonly occurs after stroke. Over 2 million Americans are currently living with chronic aphasia.[1] Anomia, or difficulty finding words despite expressing knowledge of them, is a primary symptom of aphasia that negatively impacts life participation and quality of life.[2] Improving everyday communication is the top self-identified priority of people living

with aphasia.[3] Numerous treatments for anomia exist, such as confrontation naming,[4 5] semantic-feature analysis[6 7] and phonological components analysis.[8] Compelling evidence demonstrates that naming treatments improve word retrieval, especially for trained words.[8–13] However, long-term maintenance of naming gains and generalisation of treated items to connected speech are rarely measured. Treatment-related gains are irrelevant to everyday communication if they are not durable and cannot be applied beyond the treated context. Our randomised controlled trial seeks to address this gap by evaluating the effects of a new adaptive anomia intervention based on learning and memory theories on long-term treatment maintenance and generalisation of trained words to connected speech.

Application of learning principles to anomia treatment

The Rehabilitation Treatment Specification System (RTSS)[14 15] is a framework used for guiding the development of rehabilitation interventions. As part of the RTSS framework, treatment researchers are encouraged to identify *active ingredients* and the theoretically motivated *mechanism(s) of action* underlying a treatment that are hypothesised to produce positive outcomes.[16 17] Naming treatments with hypothesised mechanisms of action based on learning principles, such as errorless or effortful learning, often incorporate active ingredients that manipulate error rate.[18]

Errorless learning paradigms typically involve providing the participant with the visual and auditory stimulus (ie, picture+name) and asking them to repeat the name of the item. This approach promotes successful learning by minimising error production which could lead to retrieval interference in future retrieval attempts. The underlying hypothesis for the principle of errorless paradigms is that treatment gains occur through Hebbian learning, that is, when a stimulus elicits a specific response the likelihood of eliciting that same response in subsequent trials increases.[19]

In contrast, effortful learning treatment paradigms typically present the visual stimulus and ask the participant to independently produce the name of the item. Error rates are typically higher when using this approach as compared with errorless learning approaches. The hypothesised mechanism of action is effortful retrieval via the testing effect[20] combined with gated Hebbian learning via performance feedback (ie, telling the patient whether their response is correct or incorrect). Effortful retrieval has been hypothesised to produce stronger memory traces leading to improved maintenance and potentially generalisation of treatment gains, while gated Hebbian learning reduces the strength of incorrect stimulus-response associations.[21–23] In the learning literature, there is robust evidence that the testing effect leads to better long-term retention of learnt material.[24–28]

Traditionally, the benefits of these two learning approaches and their corresponding underlying mechanisms have been contrasted to determine which approach is more efficacious for aphasia rehabilitation. Consistent with the learning literature, effortful retrieval has been found to increase long-term retention of trained words more successfully than errorless practice.[29] However, errorless approaches have been found to be as effective or even more effective than effortful retrieval in some studies.[30 31] A plausible explanation for these inconsistent results is: if effortful retrieval practice is too difficult, it may not provide sufficient opportunities for correct retrieval for initial acquisition, which would mitigate the otherwise clear long-term benefits of effortful retrieval. In other words, the right level of difficulty may be key for successful retention, and this level of difficulty may change over time.

Desirable difficulty in anomia treatment

Rather than contrasting errorless and effortful learning approaches, a potentially beneficial alternative is to employ the learning principle of *desirable difficulty*. This approach maximises the benefits of errorless learning for successful stimulus-response associations while also employing effortful learning to promote long-term retention. Seminal research by Bjork and Bjork[32] characterised how accuracy and effort interact to achieve desirable difficulty during learning. They drew a distinction between *retrieval strength* (ie, the ease of access to memory content given the retrieval context and recency of exposure) and *storage strength* (ie, long-term storage of and accessibility to memory representations). In their account, the act of successful memory retrieval increases the storage strength of underlying memory representations (ie, errorless retrieval). However, storage strength is moderated by the level of retrieval effort. That is, memory retrieval attempts that are effortful while still being successful lead to the strongest long-term learning effects. The current study is motivated by the premise that contextually manipulating retrieval strength to maintain desirable difficulty in anomia treatment over time will produce greater gains in storage strength, leading to more successful long-term treatment outcomes.

Desirable difficulty is a powerful framework for promoting effective language relearning in aphasia, and it can be implemented using different active ingredients. One example is *retrieval practice,* a treatment paradigm that combines errorless and effortful learning in anomia treatment. In this treatment paradigm, the stimulus name is presented auditorily, repeated by the participant and, after a time delay that allows for memory decay, is followed by retrieval attempts. This approach has proven to be superior to pure errorless learning for retention of treatment gains at 1-day and 1-week post-treatment.[4 5 29 33]

Another approach we consider to be an implementation of desirable difficulty in anomia treatment is the use of *speeded naming practice*. Conroy and colleagues[34] evaluated a speed+accuracy anomia treatment in which the naming deadline decreased from 3 s to 1 s across six treatment sessions, compared with a control condition using a traditional accuracy-only paradigm in which participants

had a 10-second naming deadline across all treatment sessions. They not only measured treatment effects for single-word naming, but they also measured the use of treated words in connected speech. They hypothesised that words that are very well learnt, especially in a speeded context, may be easier to quickly retrieve during the time pressures of connected speech. The speed+accuracy treatment produced greater gains in accuracy and speed of single word naming and had better generalisation of treated words in connected speech. The treatment gains were maintained at 1-month post-treatment. This innovative work demonstrated that speed pressure during naming treatment positively influences anomia outcomes, including generalisation to connected speech, which is a key recovery priority for people with aphasia.[3 35]

Together, these prior studies support the premise that increasing retrieval effort enhances anomia treatment outcomes. In that sense, both treatment paradigms—retrieval practice and speeded naming practice—are examples of interventions that engendered more desirable difficulty than their control conditions. However, neither of these published paradigms account for two critical considerations in aphasia rehabilitation. First, people with aphasia exhibit substantial variability in naming performance and recovery.[36] A fixed naming deadline may be too short for one person with aphasia, while it may be too long for another person due to their speed of lexical access and motor capabilities. Indeed, Evans and colleagues[37] demonstrated that people with aphasia have distinct *optimal response time cut-offs*, defined as the point at which additional response time yields diminishing gains in naming accuracy. Second, people with aphasia demonstrate intraindividual variability which impacts performance outcomes.[38] Thus, the optimal naming deadline may shift between treatment sessions; a paradigm that is optimally challenging at the start of therapy may lose its effortful quality as individuals improve throughout the course of treatment.

Adaptive treatment paradigms that calibrate difficulty based on ongoing individual performance could optimise desirable difficulty throughout the entire treatment period. One method to implement adaptive, anomia treatment is the use of patient-specific and session-specific naming deadlines. This approach to precision rehabilitation[39] could maximise the number of successful effortful practice trials per hour of treatment, minimise the number of overtly produced errors that could result in error learning and increase the speed of lexical retrieval for trained words.

Study aims and hypotheses

The aim of this randomised clinical trial (RCT; NCT05653440) is to evaluate the combined benefits of effortful and errorless learning in a novel anomia treatment that implements the learning principle of desirable difficulty. Building from prior work by our research group,[37 40–42] we implement a multinomial ex-Gaussian response time model to create a novel, effort-accuracy balanced anomia treatment that adapts based on patient-specific performance, hereafter referred to as Adaptive Balancing of Effort, Accuracy and Response Speed (Adaptive BEARS). In this treatment, we use participants' ongoing performance to establish an optimal naming deadline for each treatment session. We use a within-subject randomised, sequential crossover design to compare Adaptive BEARS to two static treatment conditions, each employing errorless and effortful treatment principles in isolation.

The primary objective of this clinical trial is to determine if Adaptive BEARS anomia treatment that uses individualised speeded naming deadlines produces superior long-term retention of treated words compared with anomia treatment with static deadlines. We hypothesise that, when compared with static treatment conditions, our novel adaptive treatment condition will produce more accurate naming of treated words at 3 months post-treatment due to increases in storage strength.

A secondary objective is to determine if Adaptive BEARS naming treatment leads to context generalisation (ie, stimulus generalisation), specifically by measuring production of treated words in connected speech during picture description tasks. We predict that, when compared with static treatment conditions, our novel adaptive treatment condition will produce more successful use of treated words in connected speech during discourse probes at 3 months post-treatment due to increases in storage strength and speed of word retrieval.

METHODS AND DESIGN

Trial design

This clinical trial uses a within-subject randomised, sequential crossover design that examines naming outcomes for three different anomia treatment conditions. Participants will complete three flashcard anomia treatments with contrasting naming deadlines in randomised order. Naming and connected speech (ie, discourse) probes will be administered twice at four distinct time points for each treatment condition: pre-treatment, 1-week post-treatment, 3-month follow-up and 6-month follow-up.

Screening for eligibility

All interested participants will complete the full Comprehensive Aphasia Test (CAT)[43] and the 175-item Philadelphia Naming Test (PNT)[44] to determine study eligibility. Additionally, interested participants will complete a brief motor speech exam with subsequent scoring of the Apraxia of Speech Rating Scale 3.5[45] to determine presence of a motor speech disorder. Data from participants with a motor speech disorder will be scored using leniency rules as described in the PNT.[44] Participants who are eligible for the study and wish to enrol will provide written informed consent (online supplemental material 1), led by the project administrator. Demographic information will be collected, including date of birth, gender identity, sex assigned at birth, race and ethnicity, education level,

languages spoken, handedness prior to stroke and handedness after stroke. Following written informed consent, we will confirm the diagnosis of a stroke via the participant's medical records.

Inclusion and exclusion criteria

Adults ages 18 and older will be included in the study if they meet the following inclusion criteria:

1. Existing diagnosis of chronic (>6 months) aphasia following a left-hemisphere ischaemic or haemorrhagic stroke.
2. Impaired performance on two of eight subtests of the CAT.[43]

Participants will be excluded if they meet any of the following criteria:

1. History of other acquired or progressive neurological disease.
2. Significant language comprehension impairments (ie, to ensure participants can understand and follow study procedures).
3. Unmanaged drug or alcohol dependence.
4. Severe diagnosed mood or behavioural disorders that require specialised mental health interventions (eg, currently active unmanaged mental health crisis, hospitalisation, neuromodulatory mental health interventions).

Sample size

Target sample size for this RCT was determined through an a priori power analysis for our primary analysis (see Statistical methods and Primary analysis section). We ran 500 generalised linear mixed-effects model simulations in R (R Core Team 2022) using the *simr* package.[46] The power analysis focused on detecting an interaction between condition and time point. The range of plausible parameter estimates for baseline performance and treatment effects were derived from Conroy *et al*[34] with random effect variances based on Evans *et al*.[37] We powered the study to detect the interaction of interest at a minimum effect size of OR=1.4; effects smaller than this, even if statistically significant, would be unlikely to represent clinically or theoretically meaningful differences. Results indicated that a sample size of 25 participants, with 40 treated items per condition, will provide power of 86.4% (95% CI 83.08% to 89.28%) to detect a significant interaction between condition and time point for the primary analysis. Given the long-term nature of the study, we planned conservatively for an attrition rate of 15% at 3-month follow-up, resulting in a recruitment goal of 30 participants to maintain power in the final analysis (figure 1). Participant recruitment began in November 2023 and study completion is anticipated in January 2028.

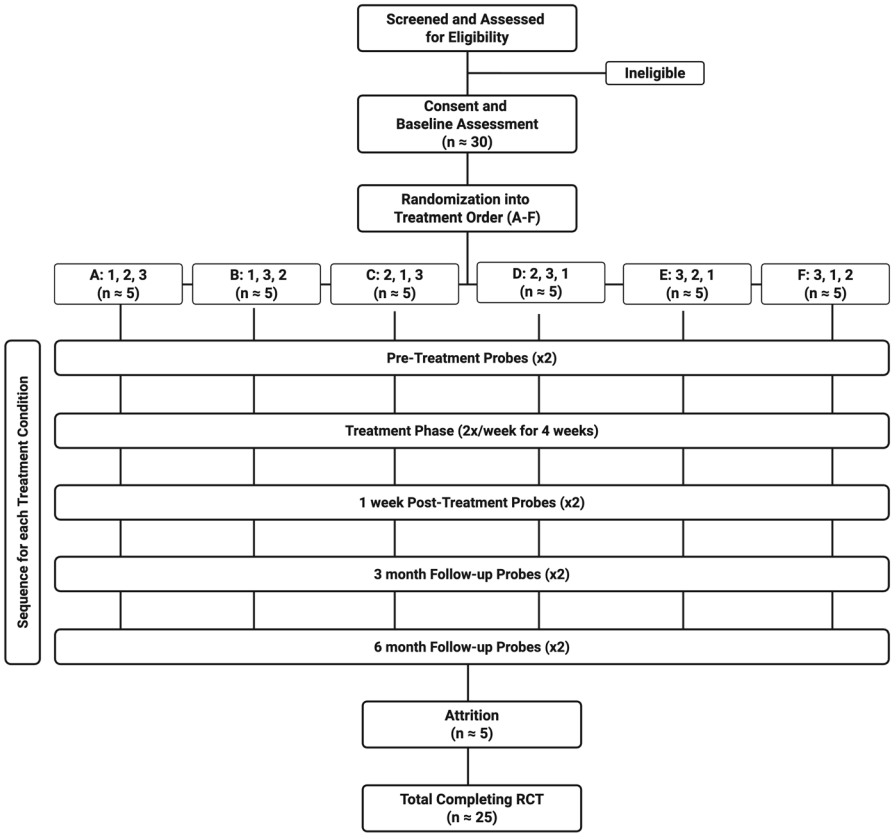

Figure 1 Flow diagram for the within-subject, sequential crossover RCT. Samples size estimates at each stage of recruitment, screening, enrolment and randomisation. Treatment order A–F represents the six counterbalanced treatment schedules for the three treatment conditions: (1) Accuracy-Maximised, (2) Effort-Maximised and (3) Adaptive BEARS. Created in BioRender. van der Stelt, C. (2026) https://BioRender.com/pznxuje. Adaptive BEARS, Adaptive Balancing of Effort, Accuracy and Response Speed; RCT, randomised clinical trial.

Assessment

Participants will complete language assessments at baseline and post-treatment to establish a comprehensive language profile for each participant and to support secondary analyses.

Baseline assessments

Following consent, baseline assessment will be performed by a certified speech-language pathologist. Language testing will include the Aphasia Bank Discourse Protocol,[47] five subtests from the Temple Assessment of Language and Short-term Memory in Aphasia,[48] picture and written versions of the Camel and Cactus Test[49] and 30 items from the Aphasia Communication Outcome Measure-Computerized Adaptive Test (CAT-ACOM).[50 51]

Post-treatment assessments

After completion of each treatment condition, the CAT-ACOM will be readministered along with the Intrinsic Motivation Inventory[52] 1-week post-treatment to measure patient-reported outcomes of communication functioning and motivation for each treatment condition, respectively. The Aphasia Bank Discourse Protocol will be readministered at the 3-month follow-up. See online supplemental appendix A for a full summary of language assessments.

Study stimuli

Noun stimuli

The pictured noun stimuli for this RCT were compiled from five sources: three existing standardised databases,[53–55] pictured nouns normed for use in a clinical trial led by author WDH (NCT02005016) and pictured nouns specifically normed for use in the current clinical trial. In some cases, a noun appeared in more than one database. In most instances, the item with the higher name agreement was retained. However, in a few instances, the item with the lower name agreement was retained because the image better reflected the target word in the USA's cultural context. Items with name agreement lower than 70% were removed. The resulting compiled battery has 862 pictured noun stimuli.

Discourse probes

The discourse probes for this study consist of 15 complex scenes and 11 text-free stories extracted from published visual search (online supplemental appendix B-1) and wordless picture books (online supplemental appendix B-2). Discourse probes will be administered as a single-page or multipage picture description task. Each discourse probe contains salient, untrained picture exemplars of treated and untreated nouns from each participant's stimuli list to evaluate use of treated words in connected speech. Salient nouns in each of these discourse probes were established through core lexicon analysis in neurotypical individuals.[56]

Stimuli and probe selection

To efficiently create balanced stimuli and probe lists, our research team developed a systematic coding algorithm via an interactive open-source web application (R Shiny App) that produces participant-specific .csv stimuli files with naming and discourse items. This web app allocates items into the three treatment conditions, designates the word type (ie, treated vs untreated items) based on frequency of occurrence in discourse stimuli and name agreement and it balances item difficulty across treatment lists at a level appropriate for each participant based on their anomia severity (ie, scaled PNT score). Item difficulty scores were derived from an established item-difficulty prediction model that uses word length, lexical frequency and age of acquisition.[57] Balancing is done using an anticlustering approach which aims to maximise high between-cluster similarity and high within-cluster heterogeneity while separating items into groups.[58] Output consists of a stratified random assignment of stimuli from our pre-existing word database to each treatment condition and word type for each participant. In making these assignments, the anticlustering approach seeks to (1) minimise the presence of unintended psycholinguistic confounders (eg, frequency effects), (2) maintain balanced difficulty across treatment conditions and (3) select appropriately difficult words for each participant to avoid baseline ceiling and post-treatment floor effects. Each stimuli list will consist of 40 treated words and 20 untreated words for each of the three treatment conditions. Stimuli between conditions are unique and do not overlap, resulting in a collective list of 180 items. For each treatment condition, at least 16 treated words and eight untreated words balanced based on the above criteria will be salient in a corresponding set of discourse probes to evaluate context generalisation. No words will be repeated across treatment conditions within a participant's list. The resulting .csv file specifies the naming and discourse probes for each study condition and assigns treated versus untreated words. This file is then uploaded into the treatment software (see AphasiaBEARS Software section).

Naming probe administration

Prior to each treatment condition, participants will be probed on their naming accuracy with the 40 treated and 20 untreated items presented in randomised order. No naming deadline will be imposed during naming probes, and participants will not receive feedback. Accuracy scoring will follow the rules of the PNT manual, resulting in a score of correct, incorrect or no response.

Discourse probe administration

Following naming probes, participants will be presented with discourse probes. Participants will be asked to describe what they see happening in detail, as if they are describing the scenes and stories to someone who cannot see the picture(s). Because participants with aphasia often require more time to find words than neurotypical

speakers, we determined time limits for discourse probe description to reduce testing burden while maintaining feasibility of completion for participants with aphasia. To do so, we first calculated the average description duration for each discourse item used by neurotypical individuals[56] and then multiplied that value by 1.5. For stories that contain multiple pages, the total description time will be split equally between pages in the story.

Treatment conditions

All participants will complete three treatment conditions with contrasting naming deadlines, comparing errorless learning and effortful learning to an adaptive condition designed to combine the benefits of both learning principles. For all treatment conditions, accuracy scoring will follow the same rules as those outlined above for the naming probes. Each trial will be preceded by an 800-millisecond fixation cross, followed by the item picture. For all conditions, the participant will be instructed to give a single verbal response in every trial and then to press the spacebar to indicate they have completed their response. Once the participant has provided a complete response per PNT rules, the clinician will log their response time and code trial accuracy (ie, correct or incorrect) via key press. If the participant is unable to provide a response before the naming deadline in the adaptive or effortful conditions, or if they indicate that they do not know in any condition, the trial will be coded as a no-response. Additionally, a prerecorded audio model of the target word will be presented along with its written form, regardless of participant accuracy, with onset of the model depending on the treatment condition. Participants will also receive visual accuracy feedback via a green 'thumbs-up' symbol provided for correct responses. For incorrect trials, the treating clinician will ask the participant to listen to and repeat the recorded auditory model 1–3 times (figure 2).

While this is a computer-based treatment, an experimenter will be present in all sessions to run the programme and ensure that participants stay engaged with this flashcard treatment across different treatment conditions. The protocol allows experimenters to provide *knowledge of results* feedback (eg, 'That was close, listen again') during treatment sessions. No further detailed feedback on productions will be provided. To maintain active participation if participants appear to become bored or frustrated, experimenters will validate feelings, offer occasional breaks and remind them of the study rationale.

Accuracy-maximised (Errorless) condition

For this condition, the picture and auditory/written models of the target word will be presented simultaneously (ie, a 0-second delay between presentation of the target and model), and participants will be asked to repeat each word. Providing the auditory model immediately is hypothesised to increase retrieval strength and thereby minimise production errors but also minimise resulting practice gains in storage strength for the target.

Effort-maximised condition

For this condition, there is a 10-second delay between presentation of the target picture and auditory/written model. Participants will be asked to take their time to name the picture accurately using only one word. The 10-second naming deadline provides ample time for participants to effortfully retrieve the target unaided,[37] which is hypothesised to produce larger gains in storage

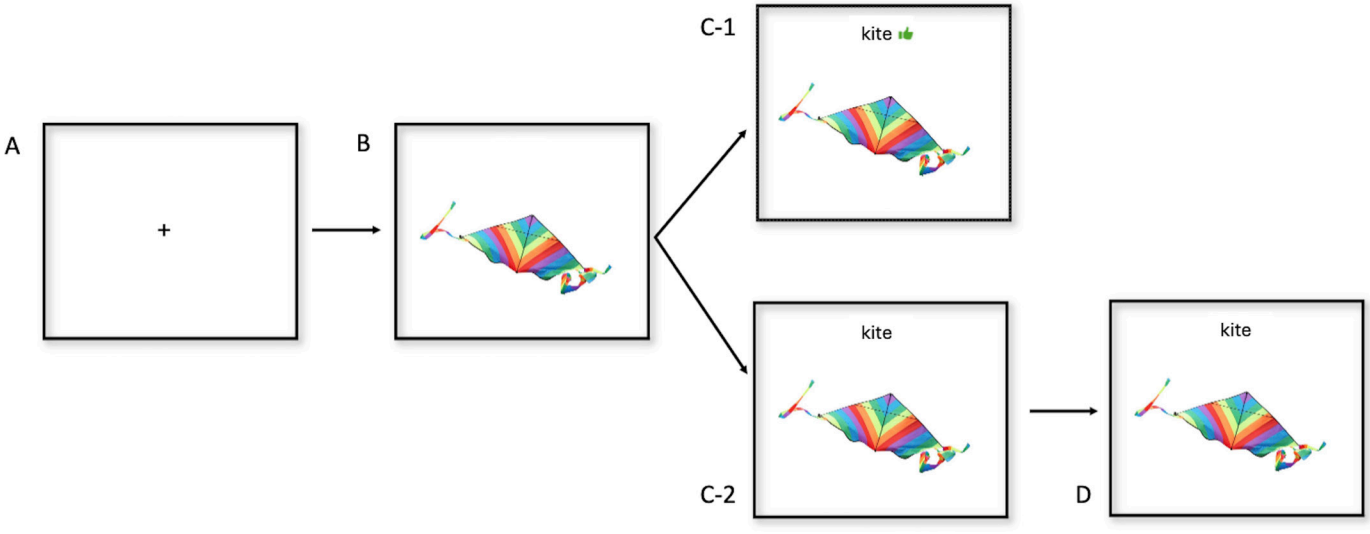

Figure 2 Illustration of treatment conditions. (A) 800 ms fixation cross appears prior to each trial. (B) Naming probe appears and participant is asked to give a single complete response before the naming deadline in each condition: Accuracy Maximised 0 s, Effort-Maximised 10 s, Adaptive BEARS 1.5–10 s. Clinician scores response time and trial accuracy. (C) Participant receives audio model of target word for all trials (C-1 and C-2) and visual feedback for correct trials (C-1). (D) Participant listens to and repeats the audio model 1–3 times for all incorrect trials. Adaptive BEARS, Adaptive Balancing of Effort, Accuracy and Response Speed.

strength on trials where participants accurately retrieve the word. However, participants will also be more likely to produce error responses in this condition, which may induce error learning. Additionally, the longer trial duration means that fewer total practice trials can be completed in this condition per session.

Adaptive BEARS condition

This novel, adaptive naming treatment condition uses a multinomial ex-Gaussian response time model[37] to determine individual speeded naming deadlines for each participant. Specifically, the model uses a multinomial distribution to estimate the proportion of correct, incorrect and no-response trials and separates ex-Gaussian distributions for correct and incorrect response times. Given this set of model parameters, we will determine the naming deadline (ie, the time when the person with aphasia hears the repetition cue) that optimises the difference between the number of correct and incorrect responses. That is, the objective being optimised is (# correct—# incorrect), and the cue timing is set to achieve the highest possible value for this objective. The goal of the optimal naming deadline is to provide ample opportunity for successful effortful retrieval while also minimising error responses, which tend to be slower than correct responses.[59] This approach is hypothesised to allow participants to provide more effortful successful retrieval attempts and fewer overt errors than either of the other conditions, leading to greater gains in storage strength for words trained in this condition. Deadlines will be adjusted for each session throughout the treatment period based on the participant's ongoing performance. The first treatment session will use a 5-second naming deadline for all participants. For each participant's second treatment session we will enter accuracy and response time data from their pretreatment naming probes and their first treatment session into the response time model to determine their individualised naming deadline. Treatment sessions 3 through 8 will then use their accuracy and response time data from the previous two treatment sessions to determine the naming deadline for the subsequent session. The model sets a minimum naming deadline duration of 1.5 s, which approximates neurotypical naming performance,[37] and a maximum duration of 10 s. To support reproducibility, see online supplemental material 2 for the R code for the ex-Gaussian response time model.

Aphasia BEARS software

We developed a web-based application, hereafter referred to as AphasiaBEARS, for study staff to administer probe and treatment sessions. AphasiaBEARS was built on Amazon Web Services,[60] using secure cloud storage to protect the integrity and confidentiality of study data. All stimuli images and audio recordings were uploaded into AphasiaBEARS at study setup. Prior to study initiation for each participant, the participant-specific .csv file (ie, see Stimuli and probe selection) will also be uploaded into the web app. Prior to starting each probe or treatment session, the research assistant has the option to test the audio signal, and enable offline fallback storage to avoid data loss such as temporary interruption in internet connectivity. Based on the participant's .csv stimuli file, AphasiaBEARS identifies the appropriate stimuli and presents them in random order. For probe sessions, items will be presented once in random order. For treatment sessions, where items will be displayed more than once, AphasiaBEARS implements randomised order without replacement, such that all 40 items will be displayed once per cycle, and no item repeats until all 40 items have been shown. The session user interface is intentionally designed to be minimalistic to promote participants' attention on the stimulus items. Sessions can be paused and resumed at any time. After completing a session, study staff will download a .zip output file that contains all session data, including a timestamped session log, coded participant responses and item-level audio recordings. See online supplemental materials 3 and 4 for example AphasiaBEARS .csv output files for an example participant. These two files can be used in conjunction with online supplemental material 2 to illustrate how the ex-Gaussian response time model uses past participant performance to establish the adaptive naming deadline for the Adaptive BEARS anomia treatment condition.

Randomisation of treatment order

After completing baseline assessment measures, participants will complete all three treatment conditions in a randomised order. The randomisation scheme was created by the trial statistician (LT) using SAS/STAT Software V.9.4. Six treatment schedules were created using counterbalancing to control order effects. The randomisation scheme was exported into Microsoft Excel and uploaded into Research Electronic Data Capture (REDCap),[61 62] the trial data management system. On enrolment, the REDCap system will randomise participants to one of the six treatment schedules. In contrast to traditional crossover designs, we do not use a washout period because stimuli are non-overlapping between treatment conditions. This approach is grounded in the well-established finding in the aphasia treatment literature that generalisation from treated to unrelated, untreated words is minimal.[8 12 31]

Study schedule

For each condition, treatment will consist of 8 sessions over 4 weeks (two per week), resulting in 24 total treatment sessions across the 3 conditions. Naming and discourse probes will be administered twice at four time points for each treatment condition: pretreatment, 1-week post-treatment, 12 weeks post-treatment±7 days (ie, 3-month follow-up) and 24 weeks post-treatment±7 days (ie, 6-month follow-up). Participation will last approximately 10 months (figure 3).

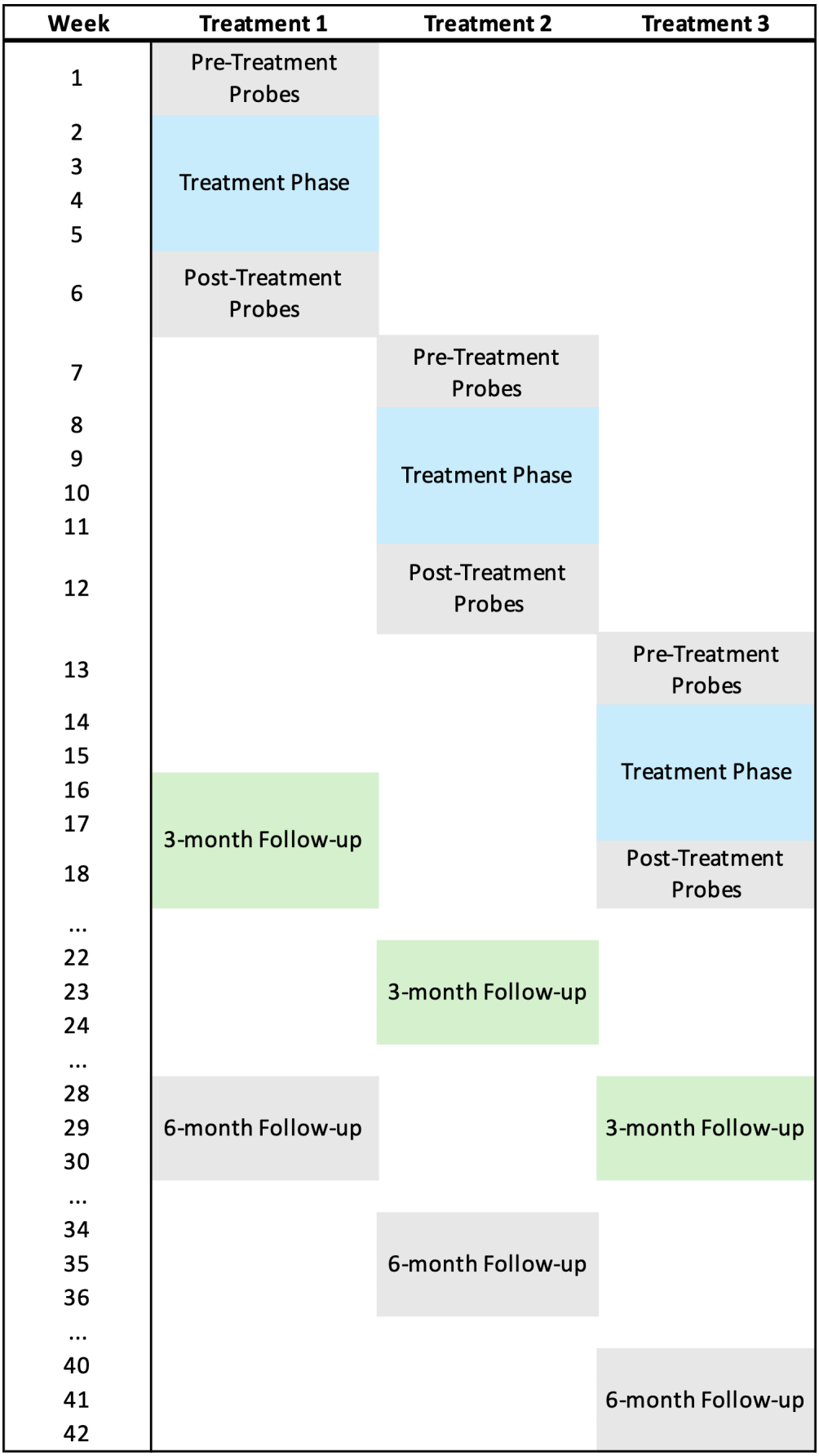

Figure 3 Study schedule for an example participant. Pretreatment probes followed by 4-week treatment phases and post-treatment testing at 1-week, 3 months (primary outcome time point) and 6 months.

Study setting

This study will be conducted in the greater Pittsburgh, Pennsylvania region. To maximise study accessibility for participants, data collection will occur at the University of Pittsburgh, in participants' homes or in a private room in a publicly available space (eg, library, community centre).

Statistical methods

Primary outcome measure

The primary outcome of this clinical trial is long-term retention of anomia treatment gains, specifically the number of treated items produced accurately on naming probes 3 months after treatment exit. We will evaluate between-condition differences across study time points (ie, pretreatment, post-treatment, 3-month follow-up, 6-month follow-up). We anticipate greater number of treated words produced accurately at 3-month follow-up compared with pretreatment. We chose to use the 3-month follow-up, rather than the 6-month follow-up, as the primary time point of interest because it reflects long-term treatment retention while being less vulnerable to potential missing data due to participant attrition.

Planned primary analysis

Following current modelling best practices,[63] we will implement item-level generalised linear mixed-effects modelling with a binomial probability distribution and logistic link function[64 65] to test for time point by treatment condition interactions using R software. The dependent variable will be item-level accuracy (1 or 0) on naming probes. Two 3-level categorical fixed effects of Condition (Adaptive BEARS, Accuracy-Maximised, Effort-Maximised) and Time point (pretreatment, post-treatment, 3-month follow-up) will be included in the model with an interaction term. For both fixed effects, we will employ treatment coding to set the Adaptive BEARS condition and the pretreatment time point as the reference levels. We will include a maximal, theoretical random effects structure for participant and item.[66 67] The effect of interest is a significant interaction between Time point (3-month follow-up>pretreatment) by Condition (Adaptive BEARS>Accuracy Maximised and Adaptive BEARS>Effort-Maximised). This finding would indicate that the Adaptive BEARS condition produces better long-term retention of naming gains. Missing data will be handled in line with best practices for clinical research.[68]

Planned secondary analyses

To evaluate generalisation of anomia treatment gains to discourse, we will test performance on discourse probes. The dependent variable will be presence or absence (1 or 0) of treated words produced during discourse probes. The fixed effect and random effect structures will mirror that of the primary analysis. The effect of interest is a significant interaction between Time point (3-month follow-up>pretreatment) by Condition (Adaptive BEARS>Accuracy Maximised and Adaptive BEARS>Effort-Maximised). This finding would indicate that the Adaptive BEARS condition produces better production of treated words in connected speech.

It is also important to distinguish treatment effects from probe exposure. Therefore, a secondary analysis will compare performance on treated naming probes to untreated control words. The dependent variable will be item-level performance on naming probes and discourse probes. Fixed effects will include Word-Type (Treated vs Untreated), Time point and an interaction term. Treatment coding will be used to set the Untreated words and pretreatment time point as reference levels. Significant differences in Word-Type would show that the treated words are associated with a significant improvement in naming performance that is not attributable to simple repeated probe exposure. We will also evaluate naming and discourse probe performance as secondary outcomes at the 1-week post-treatment and 6-month follow-up time points, aligned with the models above.

Further secondary analyses will examine change in performance in core lexicon analysis on Aphasia Bank discourse samples from initial assessment to each post-treatment follow-up time point, and explore patient-related predictors of treatment response using the Intrinsic Motivation Inventory. To characterise treatment enactment,[69] we will examine the impact of treatment on everyday communication function using patient-reported scores on the CAT-ACOM pretreatment versus post-treatment.

Data collection and participant retention

Given the study's long-term nature, participants will be permitted to take optional breaks as needed, allowing up to 18 months for completion. While participants will be instructed to refrain from participating in outside restorative anomia treatments for the duration of the study, they will not be restricted from receiving speech-language pathology services for other language modalities (eg, reading, writing, auditory comprehension) or training in alternative or compensatory expressive communication approaches. This decision was determined a priori to align with standard ethical principles in human subjects research[70] and to support participant retention by ensuring that participation in the study did not require undue restriction of clinically indicated care, while also balancing study needs for methodological rigour and internal validity. The within-subject design and randomised condition order are to mitigate potential confounds from external factors. In addition, receipt of outside therapy will be documented, examined for effects in univariate analyses and included as a covariate if statistically warranted.

Data management

Following informed consent, probe and treatment sessions will be audio and visually recorded for blinded scoring. Participant data will be de-identified using alphanumeric codes. Data will be stored on the study web application supported by Amazon Web Services.[60] Data will be

backed up on a university SharePoint server. Participant session completion will be tracked via REDcap.[61 62] Study materials with personal identifying information will be maintained for 7 years after the completion of the study and then destroyed. Participants may give consent for preservation of audio/video recordings beyond that period to use for research dissemination and educational purposes.

Data monitoring

This study does not require a Data Safety and Monitoring Board due to the minimal risk nature of behavioural research. However, the data management team will monitor data quality throughout the study during weekly meetings.

Study fidelity

Several measures will be taken to monitor study fidelity in relation to scoring probe data, treatment administration, treatment fidelity and treatment receipt.[69] The principal investigator will meet with treating study staff once per week to address questions in all these areas to ensure adherence to study protocol.

Probe scoring training

All research assistants will be trained to score probe responses following the rules outlined in the PNT. Initially, three certified speech-language pathologists (CMS, AK and MLG) scored the responses from the first three participants (see Trial status, below) and reached 90% agreement or higher. For all subsequent training, study staff will read the PNT guidelines and listen to example responses with provided accuracy codes, followed by blindly scoring participant responses. Training will be considered complete when the trainee reaches 90% or higher agreement with the certified SLP scores on at least two participants.

Treatment administration training

The treatment protocol was written by four licensed speech-language pathologists (CMS, AK, MLG and WSE) and was further edited for clarity during treatment administration of the first three participants. Then, a stepwise procedure was created to train additional study staff to implement the treatment protocol. The training sequence includes reading the treatment protocol, watching at least one prerecorded treatment session, observing at least three live treatment sessions led by trained study staff and administering at least three sessions with 100% supervision. During supervised treatment sessions, both the trainer and trainee will code participant responses, and these scores will be compared with determine readiness for independence. Study staff will be considered trained when (1) the trainee's scored response time is within 0.5 s of the trainer's scored response time, (2) the trainee's accuracy coding is at least 90% in agreement with the trainer's accuracy coding and (3) the trainee followed all treatment protocol guidelines. Each of the training steps will be supported by weekly lab meetings and individual meetings with previously trained study staff to address questions. Further clarifying edits to the treatment protocol will be added, as needed, to ensure treatment fidelity across treating study staff members.

Treatment fidelity

To maximise internal validity and reduce unintended confounding factors (eg, treatment provider drift, treatment receipt), a treatment fidelity procedure will be implemented concurrently with data collection based on treatment fidelity recommendations for RCTs.[69 71 72] Trained research assistants will review one recorded treatment session in its entirety for each treatment condition for all participants. The reviewer will note protocol deviations on a fidelity checklist that outlines the treatment steps and allowable clinician feedback in the study protocol. Deviations from the protocol will be reviewed and discussed among the study team to maximise fidelity across treating research assistants. Weekly clinical trial team meetings will be held with the principle investigator (PI) to discuss questions regarding protocol implementation to maintain high study fidelity.

Treatment receipt

Participants will be provided with task instructions at the beginning of every treatment session. A standardised script (online supplemental appendix C) will be used with every participant at the onset of treatment. Participant deviations from the protocol, noted by the treating research assistant or through the treatment fidelity procedure, will be reviewed during weekly lab meetings with the PI. Minor modifications to task instructions will be made when necessary to facilitate participant adherence to the study protocol.

Blinding

Due to the nature of the interventions, participants and treating study staff cannot be blinded to the sequence of treatment conditions. However, participants will be blinded to study hypotheses. Additionally, accuracy of naming probes and transcriptions of participant responses on discourse probes will be scored by research assistants blinded to both time point and treatment condition using audiovisual recordings.

Patient and public involvement

Patients and the public were not involved in study design or execution. A community liaison with aphasia and community aphasia support groups will be involved in participant recruitment to achieve the target sample size and in dissemination of the study.

Trial registration and protocol amendments

This trial was registered with ClinicalTrials.gov/on 7 December 2022. All protocol amendments will be approved by the Institutional Review Board at the University of Pittsburgh. Protocol amendments and deviations will be reported in subsequent publication(s) along with study findings.

ETHICS AND DISSEMINATION

The Institutional Review Board of the University of Pittsburgh approved the trial protocol (Study 21120130). All interested participants will provide written informed consent prior to study enrolment. Consent to participate in this study includes a data sharing plan. De-identified test scores, treatment and probe data, and audio and/or video recordings will be shared on Aphasia Bank (R01-DC008524) for all study participants. Aphasia Bank is a large online data repository that supports collaborative aphasia research and can be accessed online for registered users (https://aphasia.talkbank.org/).

The data sharing plan for this study also includes sharing of experimental task files, protocols, de-identified study data and statistical code on Open Science Framework (https://osf.io/) to encourage the replication, collaboration and expansion of our work by other research groups. The results of this clinical trial will be published in a peer-reviewed journal.

Privacy and confidentiality

Participant data will be de-identified using alpha-numeric codes, and data will only be accessible to study staff. Demographic information will be stored in a password-protected file.

DISCUSSION

Anomia treatments need robust evidence of long-term retention and generalisation to connected speech to support durable, functional treatment gains for people living with aphasia. This study aims to produce a novel, adaptive anomia treatment, Adaptive BEARS, that uses the learning principle of desirable difficulty to combine the benefits of effortful and errorless learning with the goal of promoting the durable re-learning and generalisation of trained words.

The Adaptive BEARS anomia treatment differs from existing anomia treatments in two key ways. First, while existing anomia treatment studies contrast learning principles, we use principles from both effortful and errorless learning to foster desirable difficulty during retrieval practice. Desirable difficulty balances accurate and effortful retrieval to maximise storage strength of learnt items.[32] Second, building on prior literature implementing speeded naming deadlines,[34 37 40–42] our computer model uses patient-specific performance to determine individualised naming deadlines that change based on ongoing patient accuracy and response time. We hypothesise that this approach will maintain desirable difficulty throughout the treatment period while maximising the number of trials completed per hour of treatment.

This RCT will evaluate the efficacy of desirable difficulty as a mechanism of action in aphasia rehabilitation and demonstrate the utility of computer modelling to individualise anomia treatment. If study predictions are supported, the Adaptive BEARS treatment will be the first empirically based anomia treatment targeting long-term retention (ie, greater than 5 weeks post treatment) for both naming gains and context generalisation to connected speech. If proven to be efficacious, Adaptive BEARS would provide an individualised approach to anomia treatment, enhancing precision rehabilitation for individuals living with aphasia.

Trial status

This trial is in active recruitment. At time of submission of this protocol manuscript, 12 participants have enrolled in the study. Three participants have completed the study. Three participants have withdrawn prior to completion. Recruitment began in November 2023 and is anticipated to end in January 2028.

Author affiliations
[1]University of Pittsburgh School of Health and Rehabilitation Sciences, Pittsburgh, Pennsylvania, USA
[2]MGH Institute of Health Professions, Charlestown Navy Yard, Massachusetts, USA
[3]VA Pittsburgh Healthcare System, Pittsburgh, Pennsylvania, USA
[4]University of Massachusetts Amherst, Amherst, MA, USA
[5]University of Pittsburgh, Pittsburgh, Pennsylvania, USA

Acknowledgements This work was presented at the 2024 Pennsylvania Speech-Language Hearing Association Conference. We are grateful to research assistants Aysha Salter-Volz, Melissa Rohm, Yukki Li, Rickea Hudson, Ashley Hartz, Naiya Caraang, Mikayla Hetrick, Maya Liesman and Megan DiCosta who assisted in stimuli compilation, norming activities, data management and/or data collection and scoring to help make this RCT possible.

Contributors CMS: writing—original draft (lead), methodology (supporting), project administration (supporting). RC: conceptualisation (supporting), methodology (supporting), funding acquisition (supporting), software (equal). WDH: conceptualisation (supporting), funding acquisition (supporting) software (supporting). JS: methodology (supporting), funding acquisition (supporting), software (equal). AK: project administration (lead) and methodology (supporting). MLG: methodology (supporting) and project administration (supporting). LT: methodology (supporting) and funding acquisition (supporting). MH: software (equal). PB: methodology (supporting), software (supporting), funding acquisition (supporting). WSE: conceptualisation (lead), methodology (lead), funding acquisition (lead), software (supporting), project administration (supporting) and writing—original draft (supporting). All authors contributed to writing—review and editing. WSE is the PI and guarantor of this manuscript. ChatGPT was used to scan the literature base for additional relevant research papers beyond what was identified by the authors. It was also used sparingly as an editing tool to improve brevity and clarity of ~8 sentences in the manuscript.

Funding Research reported in this publication was supported by the National Institute of Deafness and Other Communication Disorders of the National Institutes of Health under award numbers 1R01DC019325 and 3R01DC019325-04S1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Competing interests CMS, AK, MLG and WSE received full or partial salary support from the NIDCD grant supporting this study.

Patient and public involvement Patients and/or the public were involved in the design, or conduct, or reporting, or dissemination plans of this research. Refer to the Methods section for further details.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

ORCID iDs
Candace M van der Stelt https://orcid.org/0000-0003-1265-352X
Robert Cavanaugh https://orcid.org/0000-0002-2114-6565
William D Hula https://orcid.org/0000-0003-1635-0753
Alyssa Kelly https://orcid.org/0009-0004-6851-6449
Mara L Goodman https://orcid.org/0009-0008-0950-7496
Lauren Terhorst https://orcid.org/0000-0002-2326-8005
Mohammad Hassany https://orcid.org/0009-0004-8893-8454
Peter Brusilovsky https://orcid.org/0000-0002-1902-1464
William S Evans https://orcid.org/0000-0001-5124-3473

REFERENCES

1 Simmons-Mackie N. Aphasia in North America. Aphasia Access, 2018.
2 Hilari K, Needle JJ, Harrison KL. What are the important factors in health-related quality of life for people with aphasia? A systematic review. *Arch Phys Med Rehabil* 2012;93:S86–95.
3 Wallace SJ, Worrall L, Rose T, *et al*. Which outcomes are most important to people with aphasia and their families? an international nominal group technique study framed within the ICF. *Disabil Rehabil* 2017;39:1364–79.
4 Middleton EL, Schwartz MF, Rawson KA, *et al*. Towards a Theory of Learning for Naming Rehabilitation: Retrieval Practice and Spacing Effects. *J Speech Lang Hear Res* 2016;59:1111–22.
5 Schuchard J, Middleton EL. The Roles of Retrieval Practice Versus Errorless Learning in Strengthening Lexical Access in Aphasia. *J Speech Lang Hear Res* 2018;61:1700–17.
6 Massaro M, Tompkins CA. n.d. Feature analysis for treatment of communication disorders in traumatically brain-injured patients: An efficacy study. *Clinical Aphasiology*22:245–56.
7 Boyle M, Coelho CA. Application of Semantic Feature Analysis as a Treatment for Aphasic Dysnomia. *Am J Speech Lang Pathol* 1995;4:94–8.
8 Leonard C, Rochon E, Laird L. Treating naming impairments in aphasia: Findings from a phonological components analysis treatment. *Aphasiology* 2008;22:923–47.
9 Boyle M. Semantic Feature Analysis Treatment for Aphasic Word Retrieval Impairments: What's in a Name? *Top Stroke Rehabil* 2010;17:411–22.
10 Maddy KM, Capilouto GJ, McComas KL. The effectiveness of semantic feature analysis: an evidence-based systematic review. *Ann Phys Rehabil Med* 2014;57:254–67.
11 Oh SJ, Eom B, Park C, *et al*. Treatment Efficacy of Semantic Feature Analyses for Persons with Aphasia: Evidence from Meta-Analyses. *Commun Sci Disord* 2016;21:310–23.
12 Quique YM, Evans WS, Dickey MW. Acquisition and Generalization Responses in Aphasia Naming Treatment: A Meta-Analysis of Semantic Feature Analysis Outcomes. *Am J Speech Lang Pathol* 2019;28:230–46.
13 Yousefzade F, Memarian A, Rahimi A, *et al*. The efficacy of confrontational naming treatments for aphasia: a meta-analysis. *Aphasiology* 2025;39:942–67.
14 Hart T, Dijkers MP, Whyte J, *et al*. A Theory-Driven System for the Specification of Rehabilitation Treatments. *Arch Phys Med Rehabil* 2019;100:172–80.
15 Van Stan JH, Dijkers MP, Whyte J, *et al*. The Rehabilitation Treatment Specification System: Implications for Improvements in Research Design, Reporting, Replication, and Synthesis. *Arch Phys Med Rehabil* 2019;100:146–55.
16 Fridriksson J, Basilakos A, Boyle M, *et al*. Demystifying the Complexity of Aphasia Treatment: Application of the Rehabilitation Treatment Specification Systemx. *Arch Phys Med Rehabil* 2022;103:574–80.
17 Cherney LR, DeDe G, Hoover EL, *et al*. Applying the Rehabilitation Treatment Specification System to Functional Communication Treatment Approaches for Aphasia. *Arch Phys Med Rehabil* 2022;103:599–609.
18 Nunn K, Vallila-Rohter S, Middleton EL. Errorless, Errorful, and Retrieval Practice for Naming Treatment in Aphasia: A Scoping Review of Learning Mechanisms and Treatment Ingredients. *J Speech Lang Hear Res* 2023;66:668–87.
19 Hebb DO. Organization of Behavior. Wiley, 1949.
20 Endres T, Renkl A. Mechanisms behind the testing effect: an empirical investigation of retrieval practice in meaningful learning. *Front Psychol* 2015;6.
21 Wing EA, Marsh EJ, Cabeza R. Neural correlates of retrieval-based memory enhancement: An fMRI study of the testing effect. *Neuropsychologia* 2013;51:2360–70.
22 Ferreira CS, Charest I, Wimber M. Retrieval aids the creation of a generalised memory trace and strengthens episode-unique information. *Neuroimage* 2019;201.
23 Ye Z, Shi L, Li A, *et al*. Retrieval practice facilitates memory updating by enhancing and differentiating medial prefrontal cortex representations. *Elife* 2020;9:e57023.
24 Larsen DP, Butler AC, Roediger III HL. Repeated testing improves long-term retention relative to repeated study: a randomised controlled trial. *Med Educ (Chicago Ill)* 2009;43:1174–81.
25 Roediger HL, Karpicke JD. Test-enhanced learning: taking memory tests improves long-term retention. *Psychol Sci* 2006;17:249–55.
26 Butler AC. Repeated testing produces superior transfer of learning relative to repeated studying. *J Exp Psychol Learn Mem Cogn* 2010;36:1118–33.
27 Roediger HL, Butler AC. The critical role of retrieval practice in long-term retention. *Trends Cogn Sci (Regul Ed)* 2011;15:20–7.
28 Azzam MB, Easteal RA. Retrieval Practice for Improving Long-Term Retention in Anatomical Education: A Quasi-Experimental Study. *MedSciEduc* 2021;31:1305–10.
29 Middleton EL, Schwartz MF, Rawson KA, *et al*. Test-enhanced learning versus errorless learning in aphasia rehabilitation: testing competing psychological principles. *J Exp Psychol Learn Mem Cogn* 2015;41:1253–61.
30 Conroy P, Sage K, Lambon Ralph MA. Errorless and errorful therapy for verb and noun naming in aphasia. *Aphasiology* 2009;23:1311–37.
31 Fillingham JK, Sage K, Ralph MAL. Treatment of anomia using errorless versus errorful learning: are frontal executive skills and feedback important? *Int J Lang Commun Disord* 2005;40:505–23.
32 Bjork EL, Bjork RA. Making things hard on yourself, but in a good way: creating desirable difficulties to enhance learning.
33 Middleton EL, Schwartz MF. Errorless learning in cognitive rehabilitation: a critical review. *Neuropsychol Rehabil* 2012;22:138–68.
34 Conroy P, Sotiropoulou Drosopoulou C, Humphreys GF, *et al*. Time for a quick word? The striking benefits of training speed and accuracy of word retrieval in post-stroke aphasia. *Brain (Bacau)* 2018;141:1815–27.
35 Worrall L, Sherratt S, Rogers P, *et al*. What people with aphasia want: Their goals according to the ICF. *Aphasiology* 2011;25:309–22.
36 Knopman DS, Selnes OA, Niccum N, *et al*. Recovery of naming in aphasia: relationship to fluency, comprehension and CT findings. *Neurology (ECronicon)* 1984;34:1461–70.
37 Evans WS, Hula WD, Quique Y, *et al*. How Much Time Do People With Aphasia Need to Respond During Picture Naming? Estimating Optimal Response Time Cutoffs Using a Multinomial Ex-Gaussian Approach. *J Speech Lang Hear Res* 2020;63:599–614.
38 Duncan ES, Schmah T, Small SL. Performance Variability as a Predictor of Response to Aphasia Treatment. *Neurorehabil Neural Repair* 2016;30:876–82.
39 RELEASE Collaborators, Brady MC, Ali M, *et al*. Precision rehabilitation for aphasia by patient age, sex, aphasia severity, and time since stroke? A prespecified, systematic review-based, individual participant data, network, subgroup meta-analysis. *Int J Stroke* 2022;17:1067–77.
40 Evans WS, Quique YM. Understanding Speed-Accuracy Processing Dynamics in Aphasia Using Response Time Modeling. *Semin Speech Lang* 2021;42:240–55.
41 Evans WS, Cavanaugh R, Quique Y, *et al*. Playing With BEARS: Balancing Effort, Accuracy, and Response Speed in a Semantic Feature Verification Anomia Treatment Game. *J Speech Lang Hear Res* 2021;64:3100–26.
42 Evans WS, Hula WD, Starns JJ. Speed-Accuracy Trade-Offs and Adaptation Deficits in Aphasia: Finding the "Sweet Spot" Between Overly Cautious and Incautious Responding. *Am J Speech Lang Pathol* 2019;28:259–77.
43 Swinburn K, Porter G, Howard D. Comprehensive aphasia test. 2012. Available: https://doi.apa.org/doi/10.1037/t13733-000 [Accessed 12 Feb 2025].
44 Roach A, Schwartz MF, Martin N, *et al*. The Philadelphia Naming Test: Scoring and rationale. *Clinical Aphasiology* 1996;24:121–33.

45 Duffy JR, Martin PR, Clark HM, *et al*. The Apraxia of Speech Rating Scale: Reliability, Validity, and Utility. *Am J Speech Lang Pathol* 2023;32:469–91.
46 Green P, MacLeod CJ. Methods in ecology and evolution. In: *SIMR: an r package for power analysis of generalized linear mixed models by simulation*. 2015.
47 MacWhinney B, Fromm D, Forbes M, *et al*. AphasiaBank: Methods for studying discourse. *Aphasiology* 2011;25:1286–307.
48 Martin N, Minkina I, Kohen FP, *et al*. Assessment of linguistic and verbal short-term memory components of language abilities in aphasia. *J Neurolinguistics* 2018;48:199–225.
49 Bozeat S, Lambon Ralph MA, Patterson K, *et al*. Non-verbal semantic impairment in semantic dementia. *Neuropsychologia* 2000;38:1207–15.
50 Hula WD, Doyle PJ. The Aphasia Communication Outcome Measure: Motivation, Development, Validity Evidence, and Interpretation of Change Scores. *Semin Speech Lang* 2021;42:211–24.
51 Hula WD, Doyle PJ, Stone CA, *et al*. The Aphasia Communication Outcome Measure (ACOM): Dimensionality, Item Bank Calibration, and Initial Validation. *J Speech Lang Hear Res* 2015;58:906–19.
52 Ryan RM. Control and information in the intrapersonal sphere: An extension of cognitive evaluation theory. *J Pers Soc Psychol* 1982;43:450–61.
53 Moreno-Martínez FJ, Montoro PR. An ecological alternative to Snodgrass & Vanderwart: 360 high quality colour images with norms for seven psycholinguistic variables. *PLoS ONE* 2012;7:e37527.
54 Brodeur MB, Guérard K, Bouras M. Bank of Standardized Stimuli (BOSS) phase II: 930 new normative photos. *PLoS ONE* 2014;9:e106953.
55 Brodeur MB, Dionne-Dostie E, Montreuil T, *et al*. The Bank of Standardized Stimuli (BOSS), a new set of 480 normative photos of objects to be used as visual stimuli in cognitive research. *PLoS ONE* 2010;5:e10773.
56 Li Y. Long-term stability, test-retest reliability, and lexical properties of core nouns from novel discourse stimuli. University of Pittsburgh; 2024. Available: https://d-scholarship.pitt.edu/45934/1/Li_MS_Thesis_Final_2.pdf
57 Fergadiotis G, Kellough S, Hula WD. Item Response Theory Modeling of the Philadelphia Naming Test. *J Speech Lang Hear Res* 2015;58:865–77.
58 Papenberg M, Klau GW. Using anticlustering to partition data sets into equivalent parts. *Psychol Methods* 2021;26:161–74.
59 Starns J, Evans WS, Hula WD. Response time modeling to predict optimal person- and item-specific confrontation naming deadlines in stroke survivors with aphasia. In: *In preparation*. University of Massachusetts Amherst,
60 Amazon Web Services. Aphasia BEARS. 2023. Available: https://aphasiabears.com/signin [Accessed 17 Dec 2025].
61 Harris PA, Taylor R, Minor BL, *et al*. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* 2019;95:103208.
62 Harris PA, Taylor R, Thielke R, *et al*. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–81.
63 Meteyard L, Davies RAI. Best practice guidance for linear mixed-effects models in psychological science. *J Mem Lang* 2020;112:104092.
64 Tango T. Repeated measures design with generalized linear mixed models for randomized controlled trials. Chapman and Hall/CRC; 2017. Available: https://www.taylorfrancis.com/books/9781498747905 [Accessed 16 Dec 2025].
65 Koner S. A scientific review on advances in statistical methods for crossover design. arXiv; 2024. Available: https://arxiv.org/abs/2410.08441 [Accessed 16 Dec 2025].
66 Barr DJ, Levy R, Scheepers C, *et al*. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *J Mem Lang* 2013;68:255–78.
67 Matuschek H, Kliegl R, Vasishth S, *et al*. Balancing Type I error and power in linear mixed models. *J Mem Lang* 2017;94:305–15.
68 Heymans MW, Twisk JWR. Handling missing data in clinical research. *J Clin Epidemiol* 2022;151:185–8.
69 Behn N, Harrison M, Brady MC, *et al*. Developing, monitoring, and reporting of fidelity in aphasia trials: core recommendations from the collaboration of aphasia trialists (CATs) trials for aphasia panel. *Aphasiology* 2023;37:1733–55.
70 Satoer D, Vork J, Visch-Brink E, *et al*. Cognitive functioning early after surgery of gliomas in eloquent areas. *JNS* 2012;117:831–8.
71 Bellg AJ, Borrelli B, Resnick B, *et al*. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol* 2004;23:443–51.
72 Páez A, Nunan D, McCulloch P, *et al*. The influence of intervention fidelity on treatment effect estimates in clinical trials of complex interventions: a metaepidemiological study. *J Clin Epidemiol* 2025;177:111610.