

DISCUSSION

Designing remote synchronous auditory comprehension assessment for severely impaired individuals with aphasia

Holly Robson¹  | Harriet Thomasson¹ | Matthew H. Davis² 

¹Language and Cognition, Psychology and Language Sciences, University College London, London, UK

²MRC Cognition and Brain Sciences Unit, University of Cambridge, Cambridge, UK

Correspondence

Holly Robson, Language and Cognition, Psychology and Language Sciences, University College London, Chandler House, 2 Wakefield Street, London, WC1N 1PJ, UK.

Email: h.robson@ucl.ac.uk

Funding information

Medical Research Council, Grant/Award Number: MR/T028629/1

Abstract

Background: The use of telepractice in aphasia research and therapy is increasing in frequency. Teleassessment in aphasia has been demonstrated to be reliable. However, neuropsychological and clinical language comprehension assessments are not always readily translatable to an online environment and people with severe language comprehension or cognitive impairments have sometimes been considered to be unsuitable for teleassessment.

Aim: This project aimed to produce a battery of language comprehension teleassessments at the single word, sentence and discourse level suitable for individuals with moderate-severe language comprehension impairments.

Methods: Assessment development prioritised response consistency and clinical flexibility during testing. Teleassessments were delivered in PowerPoint over Zoom using screen sharing and remote control functions. The assessments were evaluated in 14 people with aphasia and 9 neurotypical control participants. Modifiable assessment templates are available here: <https://osf.io/r6wfm/>.

Main Contributions: People with aphasia were able to engage in language comprehension teleassessment with limited carer support. Only one assessment could not be completed for technical reasons. Statistical analysis revealed above chance performance in 141/151 completed assessments.

Conclusions: People with aphasia, including people with moderate-severe comprehension impairments, are able to engage with teleassessment. Successful teleassessment can be supported by retaining clinical flexibility and maintaining consistent task demands.

KEYWORDS

aphasia, assessment, auditory comprehension, telehealth, teleassessment

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. *International Journal of Language & Communication Disorders* published by John Wiley & Sons Ltd on behalf of Royal College of Speech and Language Therapists.



WHAT THIS PAPER ADDS

What is already known on the subject

- Teleassessment for aphasia is reliable but assessment of auditory comprehension is difficult to adapt to the online environment. There has been limited evaluation of the ability of people with severe aphasia to engage in auditory comprehension teleassessment.

What this paper adds to existing knowledge

- Auditory comprehension assessment can be adapted for videoconferencing administration while maintaining clinical flexibility to support people with severe aphasia.

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

- Teleassessment is time and cost effective and can be designed to support inclusion of severely impaired individuals.

INTRODUCTION

There are numerous advantages of telepractice for people with aphasia. Telepractice can overcome restrictions to in-person working such as barriers induced by restrictions on movement (reduced mobility, transport issues, COVID-19), geographical location and limited time (Doub et al., 2021; Rao et al., 2022). In clinical practice, telerehabilitation can increase rehabilitation dosage, unlock access to specialist services and reduce health inequalities (Khairat et al., 2019; Weidner & Lowman, 2020). Benefits for research include improving access to rare clinical profiles, thereby increasing and diversifying samples (Koonin et al., 2020). Further practical advantages include reduction in travel time and associated costs for researchers and participants and greater ease of scheduling and video/auditory recording during data collection.

Research into teleassessment conducted online via videoconferencing in aphasia and other acquired language impairments has demonstrated overall good fidelity and high reliability when compared with standard in-person assessment (Choi et al., 2015; Guo et al., 2017; Hall et al., 2013). Modification of the Western Aphasia Battery—Revised (WAB: Kertesz, 2007) and Boston Diagnostic Aphasia Examination—Short Form (BDAE: Goodglass et al., 2001) for online use demonstrated significant group- and subject-level consistency with the strongest relationships for auditory comprehension subtests (Dekhtyar et al., 2020; Hill et al., 2009). Teleassessment also has generally high levels of acceptability (Altaib & Meteyard,

2023) with one study indicating that only 15% of individuals preferred in-person assessment (Dekhtyar et al., 2020).

Many researchers stress, however, that there is a proportion of individuals for whom telepractice is not appropriate. Lack of access to hardware and internet connectivity are barriers that can be overcome for research (see Fleming et al. (2020) for an example) but are more problematic in clinical practice. Patient-related factors include significant perceptual, language comprehension and cognitive impairment (Doub et al., 2021). For example, individuals with traumatic brain injury and reduced attention were found to perform more poorly during telerehabilitation than in-person rehabilitation (Georgeadis et al., 2004). Other studies have had low representation of individuals with comprehension impairments. In Dekhtyar et al.'s (2020) WAB teleassessment study, only 5%–10% of participants were observed to have a comprehension impairment at the single word level. In Hill et al.'s (2009) examination of the BDAE, the mean single word comprehension score was 14/16 for the most severe aphasia group. Such under-representation of severe aphasia is reflective of the literature more generally (as discussed in Murray et al., 2018). This is a disappointing pattern given these individuals also demonstrate poorer therapeutic outcomes (Paolucci et al., 2005). As such, it is important that researchers make efforts to include these individuals in research and to make research studies accessible. The number of research studies employing remote data collection techniques is likely to increase and care must be taken in designing remote assessment and

experimental paradigms suitable for a wide range of participants' abilities.

Obtaining language comprehension and cognition data over videoconferencing is more problematic than eliciting language production data (although the opposite pattern occurs for scoring). This is because traditional neuropsychology comprehension and cognition assessments typically require a pointing response to pictures or objects. Researchers have overcome these issues by mailing objects to participants and requesting the camera to be re-positioned in order to see pointing responses (Rao et al., 2022); 'pointing' with the cursor (Dekhtyar et al., 2020), asking participants to verbally respond by identifying a letter or number associated with a picture (Rao et al., 2022), or, historically, building custom software (Guo et al., 2017; Hill et al., 2009). These adaptations often require the support of a carer or partner alongside the participant (Rao et al., 2022) which can result in unintentional bias. Verbal responses can be unreliable in stroke aphasia (note that Rao et al., 2022 investigated individuals with primary progressive aphasia) and cursor pointing responses may require a considerable degree of fine motor control, depending on the number of distractors, and could result in selection errors, as visual feedback between participant and researcher is relatively limited.

Numerous solutions now exist for online data collection with objective responses and automated scoring such as Gorilla (<http://www.gorilla.sc/>) Inquisit (<https://www.millisecond.com/>). Such platforms are user friendly and easy to programme by researchers without extensive coding skills. Experiments can be accessed and run independently by participants and may be suitable for participants with mild impairments and familiarity with technology. However, gaining reliable research data from a severely impaired population requires clinical communication strategies and flexibility. For example, researchers may need a multimodal approach to delivering instructions, supporting auditory comprehension with writing, drawing, or by physically moving items/objects. Some participants may require repetition of training items and researchers may need to move through tests/experiments flexibly to respond to factors influencing performance, for example, participant fatigue, stress/frustration, suspected perseveration or frequent self-corrections. These issues are magnified when completing a battery of tests with varying response demands which requires repeated cognitive switching from participants. The online environment strips the researcher of many of these clinical communication strategies and hence automated solutions should be used only with participants who can effortlessly comply with task instructions. For more impaired individuals, and if task adaptations are not made, it is more

probable that the data will be influenced by wider cognitive and physical factors, masking abilities at the cognitive skill of interest and leading to data unreflective of true performance.

There are further aspects that should also be considered when undertaking remote testing using auditory stimuli. Speech comprehension assessments usually involve items being spoken to the participant by the researcher. During in-person assessment, the researcher can adapt to participants' hearing status, for example, speaking to the better ear, lowering their speaking pitch or speech rate, speaking more loudly and so on. They can also evaluate the suitability of the auditory environment to ensure optimal audibility of task materials, for example, moving to a quiet space if necessary and minimising background noise. Researchers are unable to calibrate their speech and listening conditions as successfully in an online environment; they may be unable to hear background noise, or find it harder to judge what the participant can or cannot hear. Therefore, using prerecorded auditory stimuli is recommended, and we encourage researchers to ensure that the amplitude of all recorded stimuli is normalized to a consistent, maximal level (this is possible using standard audio editing software, e.g., <https://www.audacityteam.org>). This allows participants to set the volume to an appropriate level at the start of the testing session and to leave this setting unchanged throughout testing. Once the volume is calibrated it can be set at the beginning of each subsequent testing session to ensure consistency across tests and sessions. The quality of speakers/headphones should also be considered; not all participants will necessarily have access to suitable headphones and microphones for assessment purposes. Furthermore, it may be preferable to standardise the equipment used across participants to reduce the possibility of hardware-related differences having an impact on task performance for different individuals. Finally, the option for item repetition should ideally be built into stimulus presentation (in the PowerPoint-based method that we describe herein this can be simply achieved by the researcher clicking a button to replay the sound) to account for unexpected background noise or internet dropout which can affect the reliability to comprehension assessment (Altaib & Meteyard, 2023).

AIM

We aimed to produce or modify a battery of auditory comprehension assessments to be delivered over videoconferencing. The assessment battery was designed to facilitate involvement of individuals with impaired comprehension and executive functioning.

METHODS

Design principles

An assessment battery was designed adhering to the following principles:

1. Participants should be able to complete assessments independently without the help of a carer or study partner.
2. Visual, motor and executive demands should be minimised.
3. Consistency of motor and executive demands between tests should be prioritised by ensuring consistency in response requirements.
4. Auditory stimuli should be standardised (i.e., prerecorded) and embedded into testing materials but with flexibility to repeat items as required.
5. Researchers should be able to control when participants move between trials and between tasks so as to retain the flexibility of in-person testing and be sensitive to participant needs, for example, fatigue, frustration and attention status.
6. Task responses should be clearly visible to researcher and research participant.
7. Task delivery should be compatible with a wide variety of home computing devices (including laptop and desktop computers and tablets).

Assessment battery

Auditory comprehension was assessed at the discourse, sentence and single word level with a range of standard clinical and custom, in-house tasks:

1. Modified version of the Discourse Comprehension Test (Brookshire & Nicholas, 1993). Following MacKenzie (2000), discourse vocabulary was modified to be suitable for British participants. A subset of five test stories and one practice story was selected to be the most culturally appropriate.
2. Semantic probe items from the BDAE extended subtests (Goodglass et al., 2001).
3. WAB yes/no auditory comprehension questions. A subset of 13/20 questions were selected to be appropriate for use in an online environment.
4. An in-house sentence verification test. Participants were required to decide whether the final word in a spoken sentence was congruent or incongruent.
5. In-house single word-picture verification test. Participants were required to decide whether a spoken word matched a picture.

Semantic processing was further assessed using written synonym judgement and phonological input processing assessed using nonword discrimination tasks (Psychological Assessment of Language Processing Abilities [PALPA] 50 and PALPA 1, respectively: Kay et al., 1992).

Assessment materials

All tests were delivered in PowerPoint. PowerPoint was selected to best replicate in-person neuropsychological testing. It allowed a single test item to be presented per slide, therefore maintaining focus on the trial item. All tests in the assessment battery used a two-alternative forced-choice response, with on-screen buttons indicating either yes/no or same/different responses. Large response boxes were created on each PowerPoint slide using the animation function. When used in conjunction with screen-sharing options and remote control options in video conference software, these response buttons allowed participants to select and modify their answer if necessary and gave clear visual feedback as to which answer they had selected (Figure 1). Responses were given with a touch screen tap or mouse/trackpad click. Response cards were sent to participants who were unable to engage with a touch screen or mouse. Participants were able to hold these up to give a response which the researcher then entered into the PowerPoint slide to provide visual feedback to the participant.

Sound files were prerecorded in a sound-attenuated booth, segmented in Praat (Boersma & Weenink, 2021) and the root mean square amplitude adjusted to 70 dB. Sound files were embedded within PowerPoint files and could be played or repeated by the researcher using the arrow keys on the researcher's computer keyboard. A Visual Basic macro was used to insert the audio files into tests with a large number of items.

Data collection preparation

Participants were posted a pair of Sennheiser SC 165 USB headphones with a noise cancelling microphone (<https://www.sennheiser.com>) for use during data collection. Prior to data collection, a volume calibration task and response button training was completed. For the volume calibration, participants listened to a string of nonwords—recorded and processed in the same manner as the experimental materials—and were asked to adjust the volume on their device to a comfortable listening level using the volume control buttons on the headphones, their mouse or tablet controls. The researcher then made a note of the volume and asked the participant to set their volume

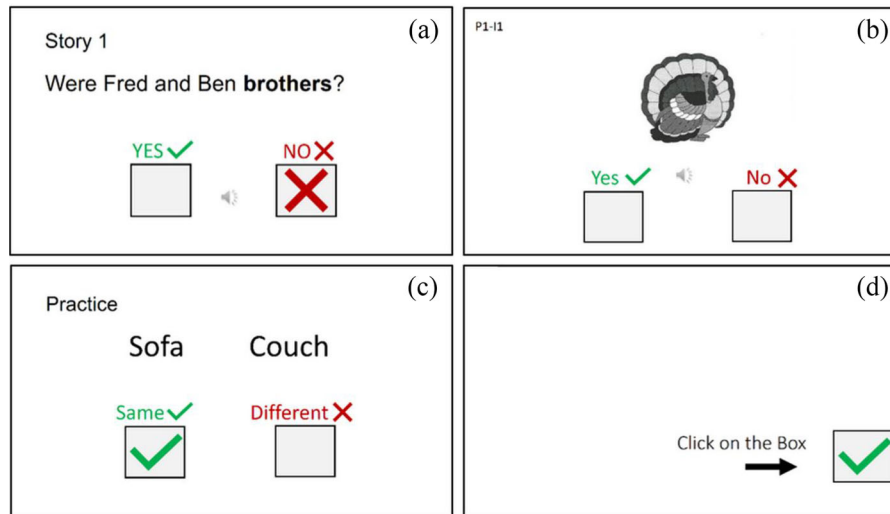


FIGURE 1 Example PowerPoint slides. (a) An example trial from a modified version of the Discourse Comprehension Test; (b) An example trial from the BDAE semantic probe subtest. (c) A practice trial for the PALPA written synonym judgement. (d) A slide from the response button training. Abbreviations: BDAE, Boston Diagnostic Aphasia Examination; PALPA, Psycholinguistic Assessment of Language Processing Abilities. [Colour figure can be viewed at wileyonlinelibrary.com]

controls to the same level in all subsequent testing sessions. For response button training, the researcher used a further PowerPoint presentation (see Figure 1d) with practice response buttons and example items. The test response buttons were positioned in the corners of the screen to ensure that participants could see the full slide. These checks also enabled the researchers to test whether participants' hardware and internet connectivity were sufficient for data collection. Best practice guidelines for environmental set up for videoconferencing data collection have been published elsewhere (e.g., Doub et al., 2021).

Videoconferencing procedure

Data collection materials were suitable for use with Microsoft Teams (<https://www.microsoft.com>) or Zoom (<https://zoom.us/>) software. We expect that other video conferencing software (e.g. FaceTime, Google Meet, Webex, etc.) could also be used as long as they provide for shared screen displays and remote control of the researcher's computer by participants. All but one participant chose Zoom due to familiarity with the software. Zoom meeting settings were set to minimise the need for participants to alter settings, for example, unmuted, video on, and to increase security, for example, admitted via a waiting room. Data collection sessions were recorded following written and verbal consent from the participant. Data were collected by sharing the PowerPoint window with the participant. The slideshow settings were changed

on the researcher's computer so that the slideshow was windowed, rather than full screen, when in presentation mode. This enabled the researcher to continue to see the participant during data collection. Participants were given remote control of the researcher's computer to enable them to enter their responses by clicking or touching the screen. However, the researcher was also able to retain control of the test session through use of the keyboard on their computer. Participants were requested to respond to the first presentation of the audio file unless stimulus presentation was disrupted by background noise or affected by internet dropout. Participants' responses were recorded manually and data recording accuracy was checked through video replay.

Example materials

Response button training, volume calibration and modifiable PowerPoint templates are available for download at <https://osf.io/r6wfm/>. Instructions for inserting audio files using a macro are also available. The Wiki on this site contains detailed instructions for running experiments in Zoom.

Ethical approval

Ethical approval for data collection was obtained from the University College London Language and Cognition Research Ethics Committee (LCD-2021-02).

Participants and preliminary data collection

Twenty-three participants had been recruited and undertaken the assessment battery at the time of manuscript preparation: $n = 9$ neurotypical control participants; $n = 6$ Wernicke's aphasia; $n = 3$ global aphasia; $n = 1$ Broca's aphasia; $n = 3$ anomic aphasia and $n = 1$ transcortical sensory aphasia. Aphasia participants were diagnosed with a modified BDAE-Short Form (Goodglass et al., 2001) presented in PowerPoint with the same response buttons as used in the neuropsychological and experimental tests with the exception of the 'commands' auditory comprehension subtest which required an action response. Participants were asked to qualitatively report on their vision and hearing. Table 1 presents participant demographics, self-reported perception and BDAE percentiles. Participants were given a £20 voucher to thank them for their participation. Participants were tested on the assessment battery using the video conferencing procedure previously described. Consent and data collection with participants with aphasia occurred over four-five 1-h sessions and over two-three 1-1/2-h sessions with neurotypical control participants.

RESULTS

All but two aphasia participants were able to complete the testing without the constant presence of a carer. These two participants required a carer to be present to further support comprehension of task instructions. For these participants, researchers sent an additional headset and a headphone splitter and carers were instructed to put their headsets on during task training and were asked to use pointing or writing to support comprehension. During presentation of test items they were instructed to take off their headphones to reduce the possibility of them unintentionally biasing participant responses. Nine participants with aphasia were supported by a carer when starting the initial videoconferencing session. This may have been for more than technical reasons, for example, desire to meet the researcher and understand the project as only five participants had support at the beginning of subsequent test sessions. Two participants with aphasia completed data collection using an iPad (<https://www.apple.com>), one used a desktop computer and the remaining participants used a laptop computer. Two participants did not have access to suitable equipment—one participant was sent a tablet holder to keep their tablet in an upright position so that their hands were free to respond and one participant was sent a Windows laptop prior to the first testing session. Three participants had significant right

arm weakness and responded using their left hand with either a touch screen or mouse. Only 1/161 assessments could not be completed for technical reasons, and 10/161 tests were not completed due to participant factors, for example, a discontinuation request and participant health. All but one participant with aphasia were able to interact with the response buttons. The remaining participant used the response cards to provide answers.

Task performance

An initial evaluation of the participants' capacity to engage with the remote assessment battery was performed by comparing participant performance to the expected chance-level performance range on each task. Chance performance may arise from a number of factors: the assessment may exceed the participant's language capacity, the participant might not comprehend the task, or they might be unable to comply with the task requirements. Binomial tests were performed in Excel using the `binom.dist` function and used to calculate chance performance for each test. The binomial test can be used to test the null hypothesis that a score obtained on an assessment falls within an expected distribution based on the assumed probability of success (Ferron & Joo, 2018). All the assessments developed/modified for this work had a yes/no or same/different response and, therefore, each trial has a 50% probability of success. The binomial test was used to identify the likelihood that the overall score on a test could occur due to chance responding. Total assessment scores with a probability of occurrence of <0.05 under this null hypothesis were interpreted as indicating above chance performance. Table 2 presents details of the number of items in each assessment, the score at which the null hypothesis can be rejected (above chance performance) and the performance range for the control participants and participants with aphasia. The participants with aphasia have been divided into those with more and less severe language comprehension impairments based on their aphasia classifications. Across the case series, there were 10 individual assessment results for which we failed to reject the null hypothesis of chance performance. Specifically, one participant with Wernicke's aphasia performed non-significantly differently from chance on spoken word verification, two participants with global aphasia did not perform differently from chance on WAB yes/no sentences, one participant with global aphasia did not perform differently from chance on nonword discrimination, one person with Wernicke's aphasia did not perform differently from chance on written synonym judgement and 5/9 participants with Wernicke's and global aphasia did not perform differently from chance on the discourse

TABLE 1 Participant demographics, perception and BDAE percentile ranks.

Pt Num.	Grp	SELF-REPORTED PERCEPTION						BDAE percentile ranks			
		Age	Vision	Hearing	Handedness	TPO	CVA type	Aphasia classification	Fluency	Aud. Comp.	Rep.
1	Aphasia	76	Historical eye surgery	No concerns	A	6y 5m	Infarction	WA	100	33	60
2	Aphasia	86	R eye squint	Mild age-related HL	R	5y 5m	Haemorrhagic	WA	100	2	5
3	Aphasia	65	Diabetes impacted	No concerns	L	7y 5m	multiple CVAs	WA/GA	70	<1	<1
4	Aphasia	54	R hemianopia	No concerns	R	7y 4m	Haemorrhagic	AA	100	63	80
6	Aphasia	65	Corrected to normal	mild HL R ear	R	13y 8m	Infarction	GA	15	18	2.5
12	Aphasia	50	Corrected to normal	No concerns	R	6y 6m	Infarction	GA	<1	9	2.5
13	Aphasia	70	R hemianopia	No concerns	L	4y 2m	Haemorrhagic	WA	100	13	2.5
17	Aphasia	62	R hemianopia	No concerns	R	14y 9m	Unknown	WA	88	10*	5
18	Aphasia	60	R hemianopia	No concerns	R	7y 5m	Haemorrhagic	WA	88	5	<1
19	Aphasia	71	Reduced PV	Mild age-related HL	R	6y 0m	Haemorrhagic	TSA	88	43	80
20	Aphasia	70	Corrected to normal	No concerns	R	7y 2m	Infarction	WA/GA	70	15	<1
21	Aphasia	58	Corrected to normal	No concerns	R	3y 5m	Arterial dissection	BA	32	67	2.5
22	Aphasia	68	No concerns	No concerns	R	10y 1m	Infarction	AA	100	100	100
23	Aphasia	74	No concerns	No concerns	L	3y 7m	Haemorrhagic	AA	88	87	90
5	Control	77	Corrected to normal	Mild age-related HL L ear	R						
7	Control	67	Corrected to normal	Mild age-related HL	R						
8	Control	67	Corrected to normal	Mild age-related HL	R						

(Continues)

TABLE 1 (Continued)

Pt Num.	Grp	Age	SELF-REPORTED PERCEPTION				BDAE percentile ranks				
			Vision	Hearing	Handedness	TPO	CVA type	Aphasia classification	Fluency	Aud. Comp.	Rep.
9	Control	75	Corrected to normal	No concerns	R						
10	Control	74	Corrected to normal	Moderate age-related HL	L						
11	Control	71	Corrected to normal	Mild HL L ear	R						
14	Control	78	Corrected to normal	Severe HF HL bilaterally	L						
15	Control	77	Corrected to normal	Mild age-related HL	L						
16	Control	83	Corrected to normal	Moderate age-related HL	R						

Abbreviations: A, ambidextrous; AA, anomia; aphasia; Aud. Comp., auditory comprehension; BA, Broca's aphasia; BDAE, Boston Diagnostic Aphasia Examination; CVA, cerebrovascular accident; GA, global aphasia; HL, hearing loss; L, left; PV, peripheral vision; R, right; Rep., repetition; TPO, time post stroke onset; TSA, transcortical sensory aphasia; WA, Wernicke's aphasia.

*Auditory Comprehension Percentile derived from Basic Word Discrimination and Complex Ideational Material only as web camera not available to assess Command responses.

TABLE 2 Neuropsychological results and chance performance analysis.

	Assessment	N test items	Minimum above chance score	Control range	WA & GA range	Other aphasia range	N at chance
1	DCT	40	25	34–39	11–33	28–34	5
2	BDAE semantic probe	60	36	57–60	36–51	50–56	0
3	WAB yes/no	13	9	13	7–11	12–13	2
4	sSPV	288	158	242–278	170–224	201–279	0
5	sWPV	105	61	92–104	52–91	81–105	1
6	PALPA 50 Synonym judgement	60	36	58–60	33–49	48–58	1
7	PALPA 1 nonword discrimination	72	36	52–72	34–63	54–68	1

Abbreviations: BDAE, Boston Diagnostic Aphasia Examination; DCT, discourse comprehension test; GA, global aphasia; N at chance, number of aphasia participants not statistically above chance on each assessment; PALPA, Psycholinguistic Assessment of Language Processing Abilities; sSPV, spoken sentence verification; sWPV, spoken word picture verification; WA, Wernicke's aphasia; WAB, Western Aphasia Battery.

comprehension test. Performance for the remaining completed assessments was greater than expected by chance (141/151).

The number of repetitions received were analysed for the spoken word picture verification task (sWPV), Boston Diagnostic Aphasia Examination Semantic Probe items, WAB yes/no questions and PALPA 1 nonword-discrimination, regardless of subsequent response accuracy. No significant difference was found between the number of repetitions provided to the groups (sWPV: Aphasia mean(SD) = 2 (3.8), Control mean(SD) = 1.25(1.8), $t_{(21)} = 0.52$, $p = 0.61$; WAB: Aphasia mean(SD) = 0.23(0.61), Control mean(SD) = 0.0(0.0), $t_{(20)} = 1.3$, $p = 0.21$; BDAE Sem: Aphasia mean(SD) = 1 (0.8), Control mean(SD) = 0.8(1.4), $t_{(20)} = 0.78$, $p = 0.44$; Nonword discrimination: Aphasia mean(SD) = 1.64(3.5), Control mean(SD) = 2.88(4.2), $t_{(20)} = -0.74$, $p = 0.47$).

Data collection problems

The two most consistent difficulties experienced by participants were problems with playing sound through the headphones (i.e., selecting the correct audio output from the options offered in the videoconferencing software) and in adjusting the computer volume. These difficulties affected control and aphasia participants in equal measure. Technical support was provided by researchers over the phone, on zoom and with training videos. Six participants were unable to communicate their volume level to the researchers. These participants therefore repeated the volume calibration step at the start of each session.

DISCUSSION

Assessment of language comprehension can be successfully adapted for synchronous testing in the online environment. Data were successfully collected from people with aphasia with and without comprehension impairments. Those with reduced comprehension were shown to be impaired at the single word level but nonetheless engaged successfully with the assessment process. Successful data collection in this project was attributed to using a relatively low-tech approach, familiar videoconferencing software (Zoom) and by ensuring flexibility in testing procedures, capacity to use clinical judgement and to ensure response consistency across/over different tasks.

Statistical analysis comparing observed performance with a chance performance baseline confirmed that participants were able to accurately engage in the tasks, indicating good comprehension of task instructions and compliance with task requirements in-line with participants' perceptual, motor and cognitive abilities. Even the most severely affected participants with aphasia performed above chance on the majority of completed assessments—only two participants were not significantly different from chance on more than one assessment. Participant 12 with global aphasia did not show above chance performance on the WAB yes/no sentences or the Discourse Comprehension Test. Participant 13 with Wernicke's aphasia did not show above chance performance on spoken word picture verification and on the Discourse Comprehension Test. Since these participants were significantly above chance performance on the other assessments, and given the similarity of stimulus and response characteristics of all tasks, it seems reasonable to conclude that chance level

performance in these participants stemmed from these tasks exceeding the participants' language comprehension capacities, rather than a more general inability to follow and adhere to task instructions in teleassessment. A further three participants did not significantly differ from chance performance at the Discourse Comprehension Test and this tallies with this assessment being the most challenging. Further consideration of the detailed pattern of results is beyond the scope of this paper which is focussed on the practicalities of teleassessment.

The aphasia group did not receive significantly more trial repetitions than the control group during assessment of spoken comprehension. Participants were instructed to give a response after the first presentation and request repetitions only because of external factors (e.g., internet dropout, background noise). However, it was not possible for the researcher to determine if this was the case as videoconferencing software cuts out extraneous background noise with good precision.

Importantly, the majority of participants were able to carry out testing without support from a carer. When a carer was required to support setup or task explanation, they were able to disengage during test items even for the least able participants. This reduces the potential for carers to inadvertently bias participants' responses. Nevertheless, that the most severely affected participants still required a carer to help setup is important. At least one of these participants was able to live independently. These participants should not be excluded from research on the basis of severity if they are able to provide informed consent with modifications and if reliable data can be elicited. The assessment and experiments described in this study can also be used for in-person testing with a one or two computer setup. The use of prerecorded stimuli is helpful in increasing consistency across testing modalities and further reduces the potential for inadvertent researcher bias. It was not the aim of the current work to evaluate test-re-test reliability between in-person and remote data collection, however, due to COVID-19 movement restrictions, it was also not possible to collect these data. Nevertheless, previous research has demonstrated good reliability of teleassessment in aphasia, including language comprehension assessment (Dekhtyar et al., 2020; Hill et al., 2009).

Previous researchers have provided best practice guidelines for setting up the environment for videoconferencing teleassessment. These guidelines include reducing potential distractions by limiting background noise, preventing interruptions and turning off other devices (Doub et al., 2021). Efforts were made to adhere to these recommendations in the current project. However, it is important to recognise that it is not always possible to maintain an

optimum environment. It is not unusual for participants to lack a desk or table and a quiet environment in which to participate due to limited space at home or large occupancy households. These participants should not be excluded from research but rather these situations reinforce the need for flexibility in assessment to allow the researcher to compensate for challenging testing environments.

Finally, and anecdotally, the researchers who collected data for this study (first and second authors) reported good task adherence, attention and engagement by participants, at a level similar to that obtained with in-person data collection. This approach had greater resource costs than in-person data collection during the development of these online tests—for example recording, segmentation and calibration of spoken stimuli. However, there was a significant cost saving during data collection with significant reductions to travel time and expenses even allowing for the costs associated with mailing equipment to participants. It is an approach that we will take again in the future and which we are pleased to recommend to the field.

Summary of clinical and research recommendations for teleassessment of auditory comprehension:

1. There should be consideration of the cognitive, sensory and motor needs of individuals with aphasia when developing and undertaking auditory comprehension assessment (see design principles).
2. The design principles laid out in this discussion paper can be applied to a wide range of clinical/neuropsychological language assessments and are not restricted to the materials used in the current project. Assessments should be selected based on clinical or research priorities.
3. Research studies requiring high levels of stimulus control and consistency should use standardised auditory stimuli and make efforts to control relevant auditory properties (e.g., volume, signal clarity) across testing sessions. However, some flexibility may be required in the approach taken to recognise the needs and capacity of individual patients.
4. The teleassessment approach laid out here is applicable to all individuals with clinically diagnosed aphasia but is unlikely to be sufficiently sensitive to identify higher-level comprehension impairments such as those observed in cognitive communication disorders.

ACKNOWLEDGEMENTS

This research was funded by a MRC Clinician Scientist Fellowship awarded to HR (MR/T028629/1)

CONFLICT OF INTEREST STATEMENT

There are no conflicts of interest to disclose.

PATIENT CONSENT STATEMENT

Informed written consent was obtained from all participants.

PERMISSION TO REPRODUCE MATERIALS FROM OTHER SOURCES

Permission to use published materials in this study was granted by the publishers.

DATA AVAILABILITY STATEMENT

Materials developed for this study are available at <https://osf.io/r6wfm/>.

ORCID

Holly Robson  <https://orcid.org/0000-0002-5391-0046>

Matthew H. Davis  <https://orcid.org/0000-0003-2239-0778>

REFERENCES

- Altaib, M.K. & Meteyard, L. (2023) A feasibility study of tele-assessment for people with aphasia in Saudi Arabia and Kuwait. *Aphasiology*, 37(8), 1198–1222. <https://doi.org/10.1080/02687038.2022.2085237>
- Boersma, P. & Weenink, D. (2021) Praat: doing phonetics by computer. Available at: <http://www.praat.org/>
- Brookshire, R.H. & Nicholas, L.E. (1993) *The Discourse comprehension test*. Tucson, AZ: Communication Skill Builders/The Psychological Corporation.
- Choi, Y.-H., Park, H.K., Ahn, K.-h., Son, Y.-j. & Paik, N.J. (2015) A telescreening tool to detect aphasia in patients with stroke. *Telemedicine and e-Health*, 21(9), 729–734.
- 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.
- Doub, A., Hittson, A. & Stark, B.C. (2021) *Conducting a virtual study with special considerations for working with persons with aphasia*. ASHA.
- Ferron, J.M. & Joo, S.-H. (2018) Binomial test. in *The SAGE encyclopedia of educational research, measurement, and evaluation*. Thousand Oaks: SAGE Publications, Inc. Available at: <https://doi.org/10.4135/9781506326139/NV-4>
- Fleming, V., Brownsett, S., Krason, A., Maegli, M.A., Coley-Fisher, H., Ong, Y.-H., Nardo, D., Leach, R., Howard, D., Robson, H., Warburton, E., Ashburner, J., Price, C.J., Crinion, J.T., Leff, A.P. (2020) Efficacy of spoken word comprehension therapy in patients with chronic aphasia: a cross-over randomised controlled trial with structural imaging. *Journal of Neurology, Neurosurgery & Psychiatry*, 92, 418–424. Available at: <https://doi.org/10.1136/jnnp-2020-324256>
- Georgeadis, A., Brennan, D., Barker, L., Baron, C. (2004) Telerehabilitation and its effect on story retelling by adults with neurogenic communication disorders. *Aphasiology*, 18(5–7), 639–652.
- Goodglass, H., Kaplan, E. & Barresi, B. (2001) *Boston diagnostic aphasia examination*, 3rd edition, (BDAE). Baltimore: Lippincott Williams & Wilkins.
- Guo, Y.E., Togher, L., Power, E., Hutomo, E., Yang, Y.-F., Tay, A., Yen, S.-C., Koh, G.C.-H., (2017) Assessment of aphasia across the International Classification of Functioning, Disability and Health using an iPad-based application. *Telemedicine and e-Health*, 23(4), 313–326.
- Hall, N., Boisvert, M. & Steele, R. (2013) Telepractice in the assessment and treatment of individuals with aphasia: a systematic review. *International Journal of Telerehabilitation*, 5(1), 27.
- Hill, A.J., Theodoros, D.G., Russell, T.G., Ward, E.C., Wootton, R. (2009) The effects of aphasia severity on the ability to assess language disorders via telerehabilitation. *Aphasiology*, 23(5), 627–642.
- Kay, J., Lesser, R. & Coltheart, M. (1992) *Psycholinguistic assessments of language processing in aphasia*. London: Psychology Press.
- Kertesz, A. (2007) *Western aphasia battery–Revised*. San Antonio, TX: PsychCorp.
- Khairat, S., Haitcoat, T., Liu, S., Zaman, T., Edson, B., Gianforcaro, R., Shyu, C.-R. (2019) Advancing health equity and access using telemedicine: a geospatial assessment. *Journal of the American Medical Informatics Association*, 26(8–9), 796–805.
- Koonin, L.M., Hoots, B., Tsang, C.A., Leroy, Z., Farris, K., Jolly, B., Antall, P., McCabe, B., Zelis, C.B.R., Tong, I., (2020) Trends in the use of telehealth during the emergence of the COVID-19 pandemic—United States. *Morbidity and Mortality Weekly Report*, 69(43), 1595.
- Mackenzie, C. (2000) The relevance of education and age in the assessment of discourse comprehension. *Clinical Linguistics & Phonetics*, 14(2), 151–161.
- Murray, L., Salis, C., Martin, N., Martin, N., Dralle, J. (2018) The use of standardised short-term and working memory tests in aphasia research: a systematic review. *Neuropsychological Rehabilitation*, 28(3), 309–351.
- Paolucci, S., Matano, A., Bragoni, M., Coiro, P., deAngelis, D., Fusco, F.R., Morelli, D., Pratesi, L., Venturiero, V., Bureca, I. (2005) Rehabilitation of left brain-damaged ischemic stroke patients: the role of comprehension language deficits. *Cerebrovascular Diseases*, 20(5), 400–406.
- Rao, L.A., Roberts, A.C., Schafer, R., Rademaker, A., Blaze, E., Esparza, M., Salley, E., Coventry, C., Weintraub, S., Mesulam, M.-M. (2022) The reliability of telepractice administration of the Western Aphasia Battery–revised in persons with primary progressive aphasia. *American Journal of Speech-Language Pathology*, 31(2), 881–895.
- Weidner, K. & Lowman, J. (2020) Telepractice for adult speech-language pathology services: a systematic review. *Perspectives of the ASHA Special Interest Groups*, 5(1), 326–338.

How to cite this article: Robson, H., Thomasson, H. & Davis, M.H. (2023) Designing remote synchronous auditory comprehension assessment for severely impaired individuals with aphasia. *International Journal of Language & Communication Disorders*, 1–11. <https://doi.org/10.1111/1460-6984.12972>