

## Research Paper

# Equivalence, modality use and patient satisfaction of telepractice administration of the scenario test in persons with primary progressive aphasia

Mirjam Gauch<sup>a,b,c,\*</sup>, Sabine Corsten<sup>a,+</sup>, Katharina Geschke<sup>c,+</sup>, Isabel Heinrich<sup>c,+</sup> and Bianca Spelter<sup>b,+</sup>

<sup>a</sup>*Faculty of Healthcare and Nursing, Catholic University of Applied Sciences Mainz, Germany*

<sup>b</sup>*Faculty of Engineering and Health, Health Campus Göttingen, University of Applied Sciences and Arts Hildesheim/Holzminden/Göttingen, Germany*

<sup>c</sup>*Department of Psychiatry and Psychotherapy University Hospital Mainz, Germany*

Received 1 November 2022

Accepted 19 May 2023

### Abstract.

**BACKGROUND:** The use of telediagnosics for people with primary progressive aphasia (PPA) could improve access to specialised care. There is a gap in research, especially regarding the evaluation of communicative-pragmatic measurement tools in a digital setting.

**OBJECTIVE:** This study examined the equivalence, modality use, and patient satisfaction of telepractice administration of the Scenario Test (ST) in people with PPA.

**METHODS:** In a cross-over design, the ST was conducted once by videoconferencing and once in person. Fifteen people with PPA participated. Participant satisfaction was assessed after each test session using a short self-designed questionnaire. The total ST scores, the use of the different communicative modalities and the participant satisfaction scores were evaluated using the equivalence and McNemar test.

**RESULTS:** Statistical equivalence was established for the present sample with regard to the total score of the ST. Regarding the use of the different modalities, no significant difference was found. Sample satisfaction was positive for both diagnostic settings, but there was no statistical equivalence of satisfaction. Severe psychiatric and cognitive symptoms affected the test performance.

**CONCLUSIONS:** This study highlights opportunities and limitations of telepractice administration of the ST in people with PPA. There is some evidence that testing with the ST via videoconferencing is feasible. Differences in the use of communication modalities and participant satisfaction should be further investigated. Influencing factors such as psychiatric symptoms and cognitive deficits should be considered in future research projects.

**Keywords:** Telepractice, Primary progressive aphasia, Communicative-pragmatic testing

\*Corresponding author: Mirjam Gauch, Faculty of Healthcare and Nursing, Catholic University of Applied Sciences Mainz, Germany. E-mail: mirjam.gauch@unimedizin-mainz.de.. ORCID ID: 0000-0001-5139-2815.

<sup>+</sup>ORCID IDs: 0000-0003-3695-8010 (S. Corsten); 0000-0001-7311-0049 (K. Geschke); 0000-0003-4310-6748 (I. Heinrich); 0000-0002-9490-6852 (B. Spelter).

## 1. Introduction

Primary progressive aphasia (PPA) is a heterogeneous, dementia syndrome (Mesulam, Wieneke, Thompson, Rogalski & Weintraub, 2012). The dominant feature is a language impairment (aphasia) that initially begins insidiously and develops a severe and global profile over the course of the disease (Tippett, 2020). PPA can be classified into non-fluent (nfvPPA), semantic (svPPA), and logopenic (lvPPA) variants (Gorno-Tempini et al., 2011). The linguistic profile of nfvPPA is characterised in particular by agrammatism in expressive speech and / or speech strain with inconsistent phonetic errors and misalignments. In svPPA, there are deficits in naming and single word comprehension. Individuals with lvPPA show word-finding disorders in spontaneous speech and naming as well as impairments in repeating sentences and phrases. Overall, the prevalence of PPA is approximately 3 : 100.000 (Coyle-Gilchrist et al., 2016; Magnin et al., 2016). However, according to Hameister, Nickels, Abel and Croot (2016) the incidence is steadily increasing. PPA usually manifests before the age of 65 and hence, earlier in life than other forms of dementia. It can impact a person's social and occupational roles (Knopman, Petersen, Edland, Cha & Roca, 2004; Mesulam et al., 2012). Due to the increasing incidence and the relatively early disease onset, addressing PPA is of great importance in both, the national economic and social context.

A comprehensive linguistic, neuropsychological, and medical examination is required to assess and diagnose PPA (Marshall et al., 2018). Based on clinical assessments and imaging techniques, PPA can be distinguished from other differential diagnoses and can be classified into the according variants. Further aims of the assessment are to determine the severity of aphasia and to identify resources and deficits while keeping the patient's burden as low as possible (Heidler, 2011).

The process of diagnosing aphasia in people with PPA should integrate all levels of the International Classification of Functioning, Disability and Health (ICF, World Health Organisation, 2013). At the structural level, imaging techniques are used to describe the extent of brain damage due to the disease. The functional level is evaluated using neurolinguistic or cognitive-oriented assessments (Peach, 2008). At the level of activity and participation, communicative-pragmatic procedures are used that indirectly or directly assess language use in everyday life (Tomp-

kins, Scott & Scharp, 2008). The Scenario Test (Van der Meulen et al., 2008) can be used to capture this level, as it measures verbal and nonverbal communication skills in severe aphasia based on everyday situations.

When conducting the Scenario Test, the examiner and the patient engage in a dialogue, whereby it is assessed how much support the examinee needs to cope with a communicative daily-life situation (van der Meulen, van de Sandt-Koenderman, Duivenvoorden & Ribbers, 2010). The test was developed and validated for individuals with poststroke aphasia and is recommended for routine use in aphasia treatment based on an international, multidisciplinary consensus (Wallace et al., 2022). To date, the Scenario Test has been validated in Dutch (van der Meulen, van de Sandt-Koenderman, Duivenvoorden & Ribbers, 2010), English (Hilari et al., 2018), and German (Nobis-Bosch et al., 2020). It is thought to be clinically useful in cases of dementia-related speech disorders (Nobis-Bosch et al., 2020) but has not been validated in this population so far.

Telehealth practices are increasingly being used in diagnostic and therapeutic processes for patients with speech disorders, as they offer advantages over face-to-face setting. This also applies to the care of people with PPA. Despite the need for detailed assessments, there are barriers to access those. Specific dementia assessments (e.g., for PPA) require high expertise and are therefore oftentimes conducted in so-called memory clinics. Due to a lack of knowledge about these facilities on the part of referring physicians and the limited capacities of institutions, not all patients with suspected dementia get access to these memory clinics (Hausner et al., 2020). The Covid-19 pandemic has revealed another potential barrier. During the first lockdown of the Covid-19 pandemic (spring 2020), face-to-face appointments at some memory clinics were partially cancelled to counteract transmission of the SARS-CoV-2 virus. To ensure the continued care of patients, telepractice was increasingly being used (Morin et al., 2021; Capra & Mattioli, 2020). However, even post COVID-19 pandemic, telepractice might provide an opportunity to conduct treatment with people in their natural environment (Cason & Cohn, 2014). Telediagnostic could save travel time and facilitate access to specialised care (Weidner & Lowman, 2020; Cotelli et al., 2017). While there is some research on telepractice for patients with PPA (Rogalski et al., 2016; Dial, Hinshelwood, Grasso, Hubbard, Gorno-Tempini & Henry, 2019), evidence on telediagnostic is lacking (Rao et al., 2022). Rao et

al. (2022) demonstrated the speech-systematic Western Aphasia Battery-Revised (WAB-R) equivalence in face-to-face and telediagnosics. So far, the transferability of communicative-pragmatic procedures like the Scenario Test to a telediagnostic setting has only been tested on very small sample sizes with people without language impairment. Therefore, the validity of the assessment is still unknown in the online setting.

Since the areas of communication and pragmatics are naturally based on face-to-face encounters (Doedens & Meteyard, 2020), it might be difficult to transfer their assessment to an online setting. Studies have shown that turn-taking and implicatures are more difficult to implement in an online setting (O'Connell, Crossley, Cammer & Morgan, 2014). Work from our own lab indicates that non-verbal communication is different and less frequently used in an online setting. For this reason, there is a particular need to examine the transferability of communicative-pragmatic approaches to a digital setting. The aim of this study was to compare face-to-face and telediagnosics using the Scenario Test in individuals with PPA regarding equivalence, modality use, and participant satisfaction.

## 2. Methods

In the following, the method of the conducted study is reported considering the TIDieR checklists (Hoffmann et al., 2014; Rhon et al., 2022).

### 2.1. Aim of the study / Research questions

This study aims to examine telediagnosics in PPA regarding the criteria of equivalence of total scores, the use of communicative modalities, and participant satisfaction for the Scenario Test. The following research questions are sought to be answered:

1. Are the total scores of the Scenario Test comparable between a face-to-face and telediagnostic setting?
2. Is the use of communicative modalities through the participants comparable between a face-to-face and telediagnostic setting?
3. Is the participant satisfaction comparable between a face-to-face and telediagnostic setting?

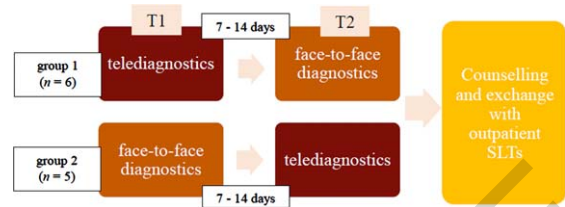


Fig. 1. Study programme. First row: Group 1 (D1-11), Second row: Group 2 (A1-10).

The hypotheses (H1-H3) of the present paper are:

H1: There are no significant differences in total Scenario Test scores in face-to-face compared to telediagnosics.

H2: There are differences in the use of communicative modalities of the participant in face-to-face compared to telediagnosics.

H3: There are no significant differences in participant satisfaction in face-to-face compared to telediagnosics.

### 2.2. Study design

The present study was a feasibility study since no study results were available on using the Scenario Test in PPA or on telediagnosics using the Scenario Test in people with aphasia at the time of study initiation. The evaluation was based on the scheme according to Dekhtyar et al. (2020), with participants undergoing face-to-face and telediagnosics in a Cross-Over-Design (see Fig. 1). At the beginning, participants were randomized into two groups. Participants of group 1 started with the telediagnostic assessment; participants of group 2 were tested face-to-face first. Participants took part in the second assessment after 7-14 days. This was considered enough time for a washout period. This procedure was chosen to balance the learning effect of the assessment. As this was a feasibility study, the sample size was calculated to be 20 participants.

### 2.3. Inclusion / Exclusion criteria

To be included in the study, participants had to have already undergone speech-systematic testing in our Memory Clinic at the time of study entry and fulfil the diagnostic criteria of PPA, according to Gorno-Tempini et al. (2011). All PPA variants were included to reflect the heterogeneity of the disorder in the study sample. All levels of PPA severity were included because the Scenario Test is suitable for mild to severe

aphasia. Since people with advanced PPA are underrepresented in studies, there is a particular need to include this group in new studies (Rao et al., 2022). Participants were required to have access to a computer and a webcam and to have support of a relative or caretaker during the teleassessment. In case potential participants did not have access to the required equipment, relatives or their outpatient speech and language therapist (SLT), who are regularly involved in the participants clinical care, were contacted and asked to provide the equipment and support for the telediagnosics for the according participant. In case this was not feasible either, participants were asked to take part in a room in the Memory Clinic where the required equipment was available while the examiner was in another room. Further inclusion criteria were an intact visual acuity (if necessary, with visual aid) and an intact hearing function (if necessary, with a hearing aid). Exclusion criteria were other neurological diseases affecting communication ability, such as stroke, Parkinson's disease or multiple sclerosis. Participants who did not communicate verbally and/or were fully dependent on the assistance of their conversational partner were excluded for ethical reasons as the test situation can be very stressful and pose a high burden to the person (Rogers & Alcaron, 1998). Information from relatives and treating outpatient SLTs served as an indicator for this classification. Participants were not required to have previous experience with the use of an online video platform (Rao et al., 2022).

#### 2.4. *Ethical aspects*

The study was conducted in the Memory Clinic of the Department of Psychiatry and Psychotherapy (University Medical Center of the Johannes Gutenberg University Mainz) from January to March 2022. Participation in the study was for all participants and did not influence further diagnosis or therapy. There was no conflict of interest for the examiner of the study. The study was conducted in accordance with the Declaration of Helsinki and approved by the Ethics Committee of the Federal Medical Association of Rhineland-Palatinate (15960 and 24.11.2021).

Persons with dementia, such as PPA, may not be able to give informed consent. Therefore, the elements of informed consent according to Beauchamp and Childress (2013) were considered. If the SLT and examiner or the doctors involved had doubts about the participant's ability to give consent, relatives and out-

patient SLTs were consulted. During the test sessions, the examiner constantly paid attention to signs of excessive demands (e.g., meta-linguistic comments, changes in facial expressions/gestures, no response). In such cases, the examiner explained again that the participant could withdraw from the study at any point and that this would not have any negative consequences for the participants.

Following the completion of both assessments, a debrief appointment could be scheduled at the request of the participant and/or their relative. The purpose of this appointment was to provide comprehensive information about the diagnostic results and to answer open questions. Further, outpatient SLTs were contacted to inform them about the diagnostic findings and to fulfil the duty of care.

#### 2.5. *Procedure*

Between December 2021 and January 2022, 37 of the Memory Clinic's patients with PPA were contacted via telephone. Following the telephone information session, written information was sent to interested patients by mail. 21 gave verbal consent to participate and were randomly assigned to two groups. Group 1 was to be initially examined in the telediagnostic setting. The participants were therefore named D1-D11. Group 2 was to receive the diagnosis in the face-to-face setting first, and the participants were consequently designated A1-A10. After randomization, the written consent form was sent in, and a physician explained the aims and benefits of the study as well as the effort associated with the study. After the medical education, the first testing took place. According to the manual of the Scenario Test, 15-45 minutes were estimated for the examination (Nobis-Bosch et al., 2020). Seven to 14 days later, the second testing was conducted.

The same unblinded SLT, referred to as "examiner", performed the test administration and evaluation. For pragmatic reasons and because validation studies of the Scenario Test show a high correlation between different assessors, this methodological weakness was accepted (Nobis-Bosch et al., 2020).

#### 2.6. *Setting*

Telediagnosics took place via the video conference software Big Blue Button (BBB; invocable GmbH, 2022), which guarantees users 100% data protection and compliance with the Basic Data Protection Regulation. In preparation for the telediag-

nostics, a link to the BBB study room was sent to the participants via e-mail. Video conferencing via BBB allowed sharing of uploaded slides while communication partners could still see each other in the video. Participants were asked to provide a self-assessment of the statements in a short self designed questionnaire. The answers were captured through the examiner. A PowerPoint presentation was used during the videoconference. This included the Scenario Test's items with a blank slide between each item to avoid participants pointing at the images, according to the Scenario Test's hand instructions. Participants and their supporting relative or SLT were asked to have a paper and a pen ready before the test to enable multimodal communication during the assessment telediagnosics. If the participant had an Augmentative and Alternative Communication (AAC) book for expressive communication, this was also provided. During the assessment, relatives were asked to always stay in the background but within reach to assist in case of technical difficulties. Following the instructions of the Scenario Test, the examiner shared the questionnaire by using the "share screen" function. Participants were then asked to answer each question.

The face-to-face testing took place in an examination room of the Memory Clinic. The room was furnished with as little stimuli as possible and was in a quiet environment so that the participants' attention was not disturbed. Participants were seated at the table opposite of the examiner and were provided with a paper and a pen. Furthermore, the test documents with the items of the Scenario Test were placed on the table.

The video camera, which documented the face-to-face and telediagnosics, was placed in a way to capture both the examiner's and participant's facial expressions and gestures during the diagnostic process.

## 2.7. Outcome measures

The Scenario Test, which measures communicative-pragmatic skills using six everyday situations represented by line drawings, was used as the primary measurement. The instructions and test administration followed the German manual instructions (Nobis-Bosch et al., 2020). There are three items per situation leading to a maximum total score of 54 points. The total score reflects how much help was needed by the participant to communicate. At the proposition level, verbal and nonverbal responses were coded. Further, the responses were

coded in terms of their modality. The test was stopped if the person with PPA did not respond to the practice items and the first scenario (Nobis-Bosch et al., 2020). Video recordings of the test situation enabled a subsequent evaluation with a simultaneous collection of quantitative and qualitative results.

As a secondary measurement tool, a short questionnaire was used to evaluate participant satisfaction with the assessment. The questionnaire comprised eight items and was self-designed based on a participant satisfaction questionnaire by Hill et al. (2009). The same questionnaire was used at both assessment conditions. If necessary, the examiner assisted in answering the questions by reading the question aloud and clarifying the principle of scale assessment in simple terms. The first item of the short questionnaire measured the affinity for technology. The following seven items measured satisfaction with the previously conducted examination. Satisfaction with the acoustic and visual presentation was evaluated. Furthermore, well-being, subjective satisfaction, willingness for further participation, comfort in the respective diagnostic setting, and satisfaction with the own performance were queried. The questionnaire used a five-point Likert scale ranging from "do not agree at all" to "fully agree". For better readability, each questionnaire item was printed on one page in a large font and landscape format. If a participant answered "do not agree at all" to continue participation, the assessment was discontinued. In the event of negative satisfaction scores after the first assessment, participants were also informed again that participation was completely voluntary and that they could withdraw from the study at any time.

## 2.8. Data analysis

To evaluate the results of the Scenario Test and short questionnaire, the data were imported into SPSS (version 27, IBM®). Both, descriptive and inferential statistics were conducted. Various forms of presentation (boxplots, pie charts, G-G charts) were created. Further, a Wilcoxon test, equivalence test, McNemar test and Kendall-Tau-b were performed. Omni Calculator (Szczepanek, 2022) was used to perform the McNemar test. In the following, the quantitative analysis procedure regarding the equivalence of the total scores, the use of modalities and participant satisfaction will be described.

For the quantitative analysis, either an intention-to-treat (evaluation in the group after randomization) or a per-protocol analysis (evaluation when the study

protocol was followed) was performed. The aim was to obtain as high external and internal validity as possible while counteracting statistical bias. Dropouts were not generally excluded from the analysis to reflect the heterogeneity of PPA in terms of severity. Instead, missing data were considered a noteworthy variable to capture the feasibility of telediagnosics in different settings and with different participants (Tabachnick & Fidell, 2013). Missing values were addressed depending on the research question and the failure process. Listwise case exclusion, pairwise case exclusion, and various imputation procedures were considered (Enders, 2010).

The primary target variables for the evaluation were the total scores in the Scenario Test and the modalities used according to the coding already described for face-to-face and telediagnosics, which were collected using the Scenario Test. The box-plot was used for the presentation. Further, various measures of location and dispersion were used for descriptive data analysis. A Q-Q plot was used to check whether the distribution of total scores conformed to a normal distribution (Field, 2018). Due to the specifics of the cross-over design, a preliminary test should be performed before inferential statistical analysis to verify that the washout period was sufficiently long and that no learning effects occurred between time point one and time point two. For this purpose, the sums of the measured values should be considered, and a test for independent samples should be used (Wellek & Blettner, 2012). In the case of a normal distribution, a t-test should be used at this point, whereas in non-parametric data, the Wilcoxon test would be indicated (Field, 2018). The test statistic used was  $p < 0.05$ .

Furthermore, an equivalence test was performed to check whether the difference between the two diagnostic settings was clinically relevant or whether the examination was equivalent in both diagnostic settings. For this purpose, the equivalence range was first defined, i.e., the range in which differences can be tolerated as clinically irrelevant (Wellek, 2002). According to the Scenario Test's manual,  $\geq 7$  performance points in the total score was designated as clinically relevant (Nobis-Bosch et al., 2020). For this reason, the equivalence range (also MCID: minimal clinical important difference) was set at  $[-7; 7]$ . Subsequently, the confidence interval was calculated to check whether it lies within the equivalence range.

The use of the different modalities was evaluated descriptively and with the help of a frequency table.

A circle diagram was used to visualise the results. The changes regarding the modalities used were additionally tested inferentially. Since the modalities used were related dichotomous variables at two test time points, the McNemar test was selected (Field, 2018). The alpha value was 0.05.

The secondary target variable was the degree of participant satisfaction, which was recorded with the aid of the short questionnaire. A bar chart was drawn to visualize the degree of agreement at item level. An equivalence test was used. A percent clinical agreement of  $> 80\%$  has been described as a high level of agreement by Hill et al. (2009). At the same time, the same authors have defined the critical difference of one scale point as a relevant difference for five-point Likert scales. For this reason, the MCID in this study is set at  $[-1; 1]$ .

### 3. Results/Findings

#### 3.1. Description of the sample

After completion of recruitment and randomisation, six participants dropped out and five participants were excluded. Dropouts occurred due to personal reasons like feeling of excessive demand, marked progression of symptoms, or an upcoming relocation. Two participants (D3, D5) had severe language impairments, so that the Scenario Test was discontinued according to the criteria described in the manual and two participants (D4, A5) withdrew from the study. One participant (D8) experienced an increase of psychiatric symptoms between the two assessment timepoints, so that he/she had to be admitted to an inpatient facility and study participation was no longer feasible. Four participants from the total sample (D3, D4, D5, A5), took part in both test sessions but showed difficulties completing the short questionnaires, resulting in missing measurements. At the end of the study (03/2022), ten complete data sets and five data sets with partially missing information were available (see Fig. 2). In the following, those participants are described who were able to take part in at least one testing and whose data were included in the statistical analysis ( $N = 15$ ). The mean age of participants was 69.27 (SD: 6.31) with a mean age of 65.27 at symptom onset (SD: 6.24). The sample consisted of 60% men and 40% women. The years of education was on average 13.67 (SD: 2.54). Of 15 participants, 33.33% had a lvPPA, 13.33% had a nfvPPA and

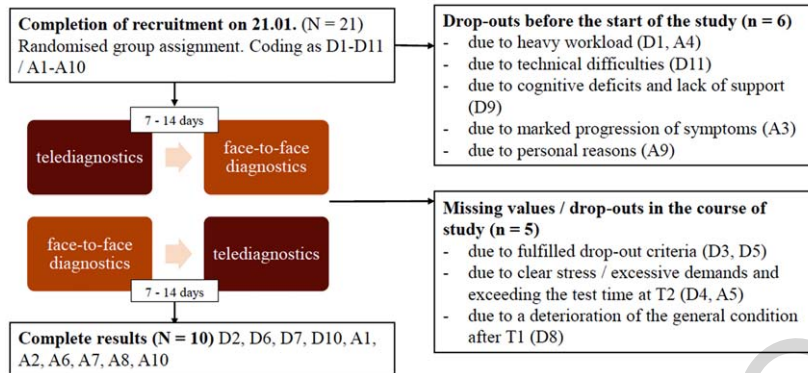


Fig. 2. Flow chart for the recruitment of the participants.

Table 1  
Demographic data. AD = Alzheimer's disease, FTLD = Frontotemporal Lobar Degeneration

Variable	M (SD) / frequencies
Age at the start of the study (years)	69.27 (6.31)
Age at symptom onset (years)	65.27 (6.24)
Gender (men / women)	9 / 6 (men / women) <sup>1</sup>
Years of education	13.67 (2.54)
PPA variant (lvPPA / nfvPPA / svPPA)	5 / 2 / 8
Pathology (AD / FTLD / unclear)	5 / 9 / 1
Depression (Yes / No)	8 / 7
Outpatient speech therapy (Yes / No)	12 / 3

<sup>1</sup> In the present sample, only the male and female gender were represented, which is why the category "diverse" is not considered in the following.

53.33% had a svPPA. Seven out of eight participants with svPPA had causal Frontotemporal Lobar Degeneration, and four out of five participants with lvPPA had underlying Alzheimer's pathology. The only participant with nfvPPA in the sample showed AD as the aetiology of PPA. In 53.3% of the sample, depression was also present. 80% of the participants were undergoing speech therapy at the start of the study. 93.3% of the participants were monolingual German speakers, one participant was an English native speaker. An overview of the demographic data of the sample is given (Table 1).

### 3.2. Evaluation of equivalence, modality use and participant satisfaction

The primary outcomes of the study participants ( $N=15$ ) are presented below (Table 2). Analysis of the primary outcome measure showed intra-individual clinically relevant differences in total

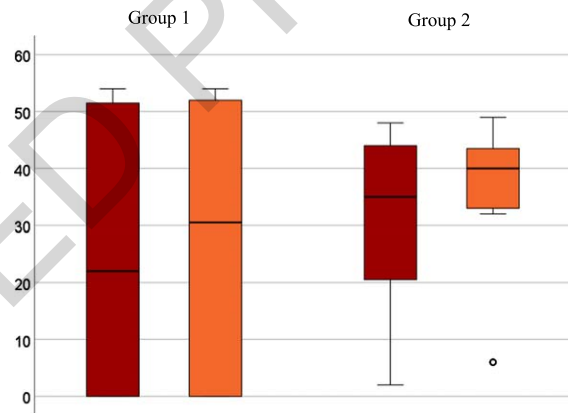


Fig. 3. Boxplot for group comparison in a telediagnostic (red) and face-to-face (orange) setting of total scores after intention-to-treat analysis ( $N=15$ ). A6 seen as outlier (o) with significant lower scores in both diagnostic settings (telediagnostic: 2; face-to-face: 6).

scores of Scenario Test ( $MCID \geq 7$  performance points) between the two test time points of six participants. The other nine did not show any clinically relevant differences. While two participants in group 1 showed ceiling effects (D6, D7), no one in group 2 reached the maximum score. Two participants in group 1 met the dropout criteria and thus achieved 0 total points. These factors led to a wide dispersion in group 1 (see Fig. 3). The Wilcoxon test reveals no significant relationship between the order of the examinations (face-to-face at time point one compared to time point 2) and the total score ( $p=0.407$ ). The confidence interval of the median calculated by bootstrapping is  $[-4.5; 0]$  and thus lies within the defined equivalence range of  $[-7; 7]$ .

Table 2

Results of the study participants (N = 15) in the scenario test, a = clinically relevant changes in total scores, b = changes in modality use, c = clinically relevant changes in participant satisfaction

	Participant Code	Total scores Scenario Test		Modality use		Affinity for technology <sup>1</sup>		Participant satisfaction (Ø)	
		digital	analogue	digital	analogue	digital	analogue	digital	analogue
group 1 (digital = t1; analogue = t2)	D2	22 <sup>a</sup>	48 <sup>a</sup>	Speaking	Speaking	-2	-2	-9 <sup>c</sup>	-6 <sup>c</sup>
	D3	0	0	Speaking	Speaking	-	-	-	-
	D4	22 <sup>a</sup>	13 <sup>a</sup>	Speaking, Gesture	Speaking, Gesture	-	-1	-	6
	D5	0	0	Speaking	Speaking	-	-	-	-
	D6	54	54	Speaking	Speaking	2	1	11	11
	D7	54	54	Speaking	Speaking	0	-1	14	14
	D8	0	-	Speaking	-	-	-	-	-
	D10	49	50	Speaking	Speaking	0	0	12 <sup>c</sup>	10 <sup>c</sup>
	A1	27 <sup>a</sup>	34 <sup>a</sup>	Speaking, Writing <sup>b</sup>	Speaking, Writing, Drawing <sup>b</sup>	2	2	9 <sup>c</sup>	10 <sup>c</sup>
	A2	48 <sup>a</sup>	40 <sup>a</sup>	Speaking <sup>b</sup>	Speaking, Writing <sup>b</sup>	0	1	10	10
group 2 (digital = t2; analogue = t1)	A5	14 <sup>a</sup>	32 <sup>a</sup>	Speaking <sup>b</sup>	Speaking, Drawing <sup>b</sup>	-	2	-	14
	A6	2	6	Speaking <sup>b</sup>	Speaking, Drawing <sup>b</sup>	-2	0	10 <sup>c</sup>	9 <sup>c</sup>
	A7	35	40	Speaking <sup>b</sup>	Speaking, Gesture <sup>b</sup>	2	1	10 <sup>c</sup>	13 <sup>c</sup>
	A8	47	47	Speaking <sup>b</sup>	Speaking, Drawing <sup>b</sup>	-1	0	9 <sup>c</sup>	11 <sup>c</sup>
	A10	41 <sup>a</sup>	49 <sup>a</sup>	Speaking <sup>b</sup>	Speaking, Writing, Gesture <sup>b</sup>	-1	-2	7 <sup>c</sup>	3 <sup>c</sup>

<sup>1</sup> According to the first item of the short questionnaire.

A more considerable variation in the use of different modalities was seen in the face-to-face, compared to telediagnostics (see Fig. 4). The preferred mode of communication in both diagnostic settings was speaking. The results of the McNemar test showed no significant differences between the use of modalities in face-to-face and telediagnostics ( $p = 0.09 - 0.85$ ).

For the evaluation of secondary outcome measures, the statement “do not agree at all” was assigned the value -2, “rather disagree” the value -1, “partly agree” the value 0, “rather agree” the value 1 and “fully agree” the value 2. Overall, the analysis showed positive satisfaction ratings for both settings, but higher satisfaction scores for the face-to-face setting (see Fig. 5). Group 2 reported higher overall satisfaction scores and showed less dispersion (face-to-face:  $M = 10$ ;  $SD = 2$ ; teledagnostic:  $M = 8.5$ ;  $SD = 2.74$ ) compared to group 1 (teledagnostic:  $M = 7$ ;  $SD = 10.74$ ; face-to-face:  $M = 7.25$ ;  $SD = 9$ ). The results of the bootstrapping show that in the present sample, no equivalence can be assumed concerning participant satisfaction in face-to-face and telediagnostics, as the confidence interval  $[-2; 0]$  exceeds the equivalence range  $[-1; 1]$ .

#### 4. Discussion

The present study answered the first research question, whether the diagnostic setting influences the total score of the Scenario Test. After statistical analysis, the hypothesis of equivalence between face-to-face and telediagnostics could be accepted. Based on the results of bootstrapping, equivalence between face-to-face and telediagnostics can be assumed regarding the total scores. Moreover, it should be taken into account that the overlapping of confidence intervals also might be a result of the small sample size. Since no significant relationship was found between the testing order and the total score, no learning effects can be assumed. Thus, it is assumed that the washout period of 7-14 days is sufficient in the present sample. Differences between groups were more prominent than the differences resulting from the diagnostic settings. This is particularly evident in the boxplot and represents a side effect of randomisation in a small sample.

The second research question, whether the diagnostic setting influences the use of communicative modalities, was partially answered. The statistical



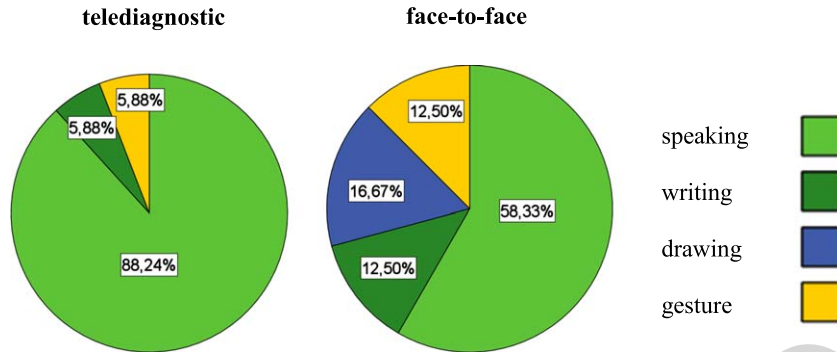


Fig. 4. Modality use in a telediagnostic (left) and face-to-face setting (right).

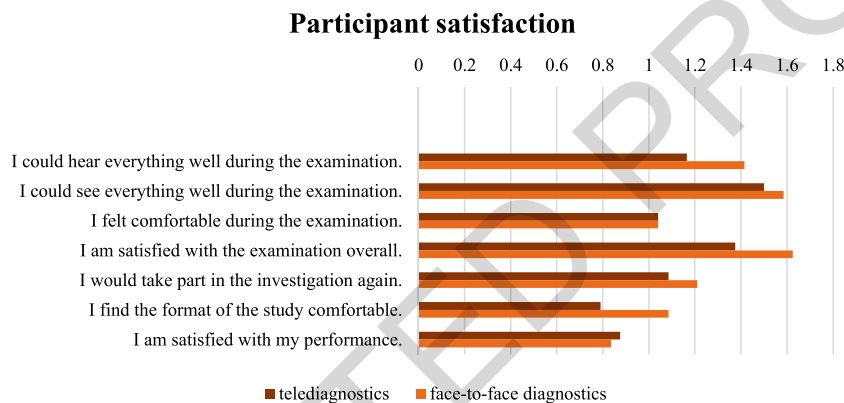


Fig. 5. Participant satisfaction with telediagnostic (red) and face-to-face (orange) setting.

evaluation and verification of the McNemar test results showed no significant differences in modality use between face-to-face and telediagnosics. Therefore, the hypothesis of differences in using communicative modalities in face-to-face and telediagnosics must be rejected. However, the increase in modality use in the face-to-face setting indicates that multimodal communication is easier to stimulate for some participants in this setting. A qualitative evaluation of these results has already been conducted and published.

(Gauch et al., 2023). The differences in modality use underline the importance of face-to-face conversation, which is also highlighted in the literature (Doedens & Meteyard, 2020). The results suggest that patients with different PPA variants benefit from multimodal communication but show differences in modality preference. During the study, most participants communicated via the modality of speaking. In some cases, access to nonverbal communication was blocked. This was the case, for example, with those participants who showed mutism. In contrast

to stroke-related mutism, the ability to compensate seems limited in PPA due to cognitive deficits. The differences between nonverbal communication in stroke-related aphasia and in PPA should be investigated in future studies and with more profound research methods. The influence of depression on communicative-pragmatic abilities might have affected performances and should also be the subject of prospective studies. The differences in modality use depending on the diagnostic setting could be reduced, for example, by technical solutions for the modalities of writing and drawing. Possibly, using 3D glasses and virtual reality could also facilitate modality use in the telediagnostic setting, as it was successfully used in former studies for people with aphasia (Giachero et al., 2020; Marshall et al., 2020). The reason why less non-verbal communication is used in the telediagnostic setting should be investigated in the future.

Furthermore, the study also answered the third research question, whether the diagnostic setting influenced participant satisfaction. For the investi-

gated sample, participant satisfaction in the different diagnostic settings is not equivalent. Even if the satisfaction of the sample was positive for both diagnostic settings, participants tended to prefer the face-to-face setting. The case of D2, the only respondent to indicate negative values in satisfaction, should also be emphasised at this point. The participant was currently in a moderate depressive episode which may have influenced his perception of the diagnostic situation and his overall scores. Overall, D2 achieved significantly higher total scores in the face-to-face setting. This may indicate that people with psychiatric symptoms (e.g., depression or delusions) benefit from direct contact with those examining or treating them. This would be consistent with a study by Tutty, Spangler, Poppleton, Ludman and Simon (2010), in which the results of cognitive-behavioural telepractice for depression lagged behind the effects of face-to-face treatments and a majority of individuals reported preferring therapy delivered in person.

#### 4.1. Future research

The test administration and evaluation by the same, unblinded person reduces the objectivity of the present work. In future studies, however, the evaluation should be carried out by two independent examiners. Further, the small sample of 15 respondents limits the significance of the statistical analysis. The implementation of larger equivalence studies is indicated to achieve higher external validity. The group of subjects was heterogeneous overall. Although the numerous drop-outs reduced the significance of the statistical results, new insights into the feasibility of telediagnosics for patients with PPA can be made. The influence of psychiatric symptoms on the feasibility of telediagnosics offers grounds for further research.

#### 4.2. Conclusion

The present study points out the opportunities and limitations of telediagnosics for people with PPA. Evidence was found that the Scenario Test in telediagnosics is feasible with PPA patients and that the results are comparable to face-to-face testing. This opens perspectives for communicative-pragmatic assessment of PPA in Dutch-, English-, and German-speaking countries. Although differences were non-significant, more communicative modalities were used in face-to-face compared to telediagnosics. This aspect should be investigated in

future studies. Regarding participant satisfaction, a basically positive attitude was recorded for both settings, which is also in line with existing evidence (Barton, Morris, Rothlind & Yaffe, 2011). The trend towards a preference for the face-to-face setting in this sample is noteworthy and should be investigated further. Nevertheless, there is a need to investigate all research questions examined in the study in a larger, representative sample of patients with PPA.

Feedback from participants indicates that people with PPA positively receive telediagnostic solutions, as shown in other studies earlier. The dropouts contributed to a reduced sample. Nevertheless, the high drop-out rate in a neurodegenerative disease is not atypical and should be seen in connection with concomitant neuropsychiatric disorders. Seeing the high potential of telediagnosics to improve care for people living with PPA, there should be more research on developing telediagnostic solutions in speech therapy. To ensure the quality of telediagnosics in PPA, further training should in future also address special features of the digital setting and the use of the procedure in people with dementia-related speech disorders.

#### Acknowledgments

This work was part of a project supported by a grant from the German Federal Ministry of Education and Research (Grant Number BMBF, 01IS19039C).

#### Declaration of interest

The authors have no conflict of interest to report. The authors alone are responsible for the content and writing of the paper.

#### References

- Barton, C., Morris, R., Rothlind, J., & Yaffe, K. (2011). Video-telemedicine in a memory disorders clinic: evaluation and management of rural elders with cognitive impairment. *Telemedicine and e-Health*, 17(10), 789-793. <https://doi.org/10.1089/tmj.2011.0083>
- Beauchamp, T. L., & Childress, J. F. (2013). *Principles of biomedical ethics*. (7<sup>th</sup> ed.). University Press.
- Capra, R., & Mattioli, F. (2020). Telehealth in neurology: an indispensable tool in the management of the SARS-CoV-2 epidemic. *Journal of Neurology*, 267(7), 1885-1886. <https://doi.org/10.1007/s00415-020-09898-x>

- Cason, J., & Cohn, E. R. (2014). Telepractice: An overview and best practices. *Perspectives on Augmentative and Alternative Communication*, 23(1), 4-17. <https://doi.org/10.1044/aac23.1.4>
- Cotelli, M., Manenti, R., Brambilla, M., Gobbi, E., Ferrari, C., Binetti, G., & Cappa, S. F. (2019). Cognitive telerehabilitation in mild cognitive impairment, Alzheimer's disease and frontotemporal dementia: a systematic review. *Journal of Telemedicine and Telecare*, 25(2), 67-79. <https://doi.org/10.1177/1357633X17740390>
- Coyle-Gilchrist, I. T., Dick, K. M., Patterson, K., Rodríguez, P. V., Wehmann, E., Wilcox, A.,... & Rowe, J. B. (2016). Prevalence, characteristics, and survival of frontotemporal lobar degeneration syndromes. *Neurology*, 86(18), 1736-1743. <https://doi.org/10.1212/WNL.0000000000002638>
- 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. [https://doi.org/10.1044/2019\\_AJSLP-19-00023](https://doi.org/10.1044/2019_AJSLP-19-00023)
- Dial, H. R., Hinshelwood, H. A., Grasso, S. M., Hubbard, H. I., Gorno-Tempini, M. L., & Henry, M. L. (2019). Investigating the utility of teletherapy in individuals with primary progressive aphasia. *Clinical Interventions in Aging*, 14, 453. <http://dx.doi.org/10.2147/CIA.S178878>
- Doedens, W. J., & Meteyard, L. (2020). Measures of functional, real-world communication for aphasia: A critical review. *Aphasiology*, 34(4), 492-514. <https://doi.org/10.1080/02687038.2019.1702848>
- Enders, C. K. (2010). *Applied missing data analysis*. Guilford press.
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics 5th ed*. Sage.
- Gauch, M., Corsten, S., Geschke, K., Heinrich, I., Leinweber, J., & Spelter, B. (2023). Differences of Modality Use between Telepractice and Face-to-Face Administration of the Scenario-Test in Persons with Dementia-Related Speech Disorder. *Brain sciences*, 13(2), 204. <https://doi.org/10.3390/brainsci13020204>
- Giachero, A., Calati, M., Pia, L., La Vista, L., Molo, M., Rugiero, C., Fornaro, C., & Marangolo, P. (2020). Conversational Therapy through Semi-Immersive Virtual Reality Environments for Language Recovery and Psychological Well-Being in Post Stroke Aphasia. *Behavioural Neurology*, 1-15. <https://doi.org/10.1155/2020/2846046>
- Gorno-Tempini, M. L., Hillis, A. E., Weintraub, S., Kertesz, A., Mendez, M., Cappa, S. F., ... & Grossman, M. (2011). Classification of primary progressive aphasia and its variants. *Neurology*, 76(11), 1006-1014. <https://doi.org/10.1212/WNL.0b013e31821103e6>
- Hausner, L., Froelich, L., von Arnim, C. A., Bohlken, J., Dodel, R., Otto, M.,... & Jessen, F. (2020). Memory clinics in Germany-structural requirements and areas of responsibility. *Der Nervenarzt*, 92(7), 708-715. <https://doi.org/10.1007/s00115-020-01007-7>
- Heidler, M. D. (2011). Einteilung, Diagnostik und Therapie von Demenzen und demenziell bedingten Sprachstörungen. *Sprache· Stimme· Gehör*, 35(02), 111-119. <https://doi.org/10.1055/s-0030-1270474>
- Hilari, K., Galante, L., Huck, A., Pritchard, M., Allen, L., & Dipper, L. (2018). Cultural adaptation and psychometric testing of The Scenario Test UK for people with aphasia. *International Journal of Language & Communication Disorders*, 53(4), 748-760. <https://doi.org/10.1111/1460-6984.12379>
- 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. <https://doi.org/10.1080/02687030801909659>
- Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D.,... & Michie, S. (2014). Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *Bmj*, 348. <https://doi.org/10.1136/bmj.g1687>
- Knopman, D. S., Petersen, R. C., Edland, S. D., Cha, R. H., & Rocca, W. A. (2004). The incidence of frontotemporal lobar degeneration in Rochester, Minnesota, 1990 through 1994. *Neurology*, 62, 506-508. <https://doi.org/10.1212/01.WNL.0000106827.39764.7E>
- Magnin, E., Démonet, J. F., Wallon, D., Dumurgier, J., Troussière, A. C., Jager, A.,... & Paquet, C. (2016). Primary progressive aphasia in the network of French Alzheimer plan memory centers. *Journal of Alzheimer's Disease*, 54(4), 1459-1471. <https://doi.org/10.3233/JAD-160536>
- Marshall, J., Devane, N., Talbot, R., Cate, A., Cruice, M., Hilari, K., MacKenzie, G., Maguire, K., Patel, A., Roper, A., & Wilson, S. (2020). A randomised trial of social support group intervention for people with aphasia: A Novel application of virtual reality. *PloS One*, 15(9), e0239715. <https://doi.org/10.1371/journal.pone.0239715>
- Marshall, C. R., Hardy, C. J., Volkmer, A., Russell, L. L., Bond, R. L., Fletcher, P. D.,... & Warren, J. D. (2018). Primary progressive aphasia: a clinical approach. *Journal of Neurology*, 265, 1474-1490. <https://doi.org/10.1007/s00415-018-8762-6>
- Mesulam, M. M., Wieneke, C., Thompson, C., Rogalski, E., & Weintraub, S. (2012). Quantitative classification of primary progressive aphasia at early and mild impairment stages. *Brain*, 135(5), 1537-1553. <https://doi.org/10.1093/brain/aww080>
- Morin, A., Pressat-Laffouilhère, T., Sarazin, M., Lagarde, J., Roue-Jagot, C., Olivieri, P.,... & Wallon, D. (2021). Telemedicine in French Memory Clinics During the COVID-19 Pandemic. *Journal of Alzheimer's Disease, (Preprint)*, 1-6. <https://doi.org/10.3233/JAD-215459>
- Nobis-Bosch, R., Bruehl, S., Krzok, F., Jakob, H., van de Sandt-Koenderman, M., & van der Meulen, I. (2020). *Szenario-Test: Testung verbaler und non-verbaler Aspekte aphasischer Kommunikation: Handbuch*. Prolog.
- O'Connell, M. E., Crossley, M., Cammer, A., Morgan, D., Allingham, W., Cheavins, B.,... & Morgan, E. (2014). Development and evaluation of a telehealth videoconferenced support group for rural spouses of individuals diagnosed with atypical early-onset dementias. *Dementia*, 13(3), 382-395.
- Peach, R. K. (2008). Global aphasia: Identification and management. In R. Chapey (Eds.) *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 586-587). Lippincott Williams & Wilkins.
- Rao, L. A., Roberts, A. C., Schafer, R., Rademaker, A., Blaze, E., Esparza, M.,... & Rogalski, E. (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. [https://doi.org/10.1044/2021\\_AJSLP-21-00150](https://doi.org/10.1044/2021_AJSLP-21-00150)

- Rhon, D. I., Fritz, J. M., Kerns, R. D., McGeary, D. D., Coleman, B. C., Farrokhi, S.,... & Hoffmann, T. (2022). TIDieR-telehealth: precision in reporting of telehealth interventions used in clinical trials-unique considerations for the Template for the Intervention Description and Replication (TIDieR) checklist. *BMC Medical Research Methodology*, 22(1), 1-11. <https://doi.org/10.1186/s12874-022-01640-7>
- Rogalski, E. J., Saxon, M., McKenna, H., Wieneke, C., Rademaker, A., Corden, M. E.,... & Khayum, B. (2016). Communication Bridge: A pilot feasibility study of Internet-based speech-language therapy for individuals with progressive aphasia. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, 2(4), 213-221. <http://dx.doi.org/10.1016/j.trci.2016.08.005>
- Rogers, M. A., & Alarcon, N. B. (1998) Dissolution of spoken language in primary progressive aphasia, *Aphasiology*, 12(7-8), 635-650. <https://doi.org/10.1080/02687039808249563>
- Tabachnick, B.G., & Fidell, L.S. (2013). *Using Multivariate Statistics 6<sup>th</sup> ed.* Pearson.
- Tippett, D. C. (2020). Classification of primary progressive aphasia: challenges and complexities. *F1000Research*, 9, 64. <https://doi.org/10.12688/f1000research.21184.1>
- Tutty, S., Spangler, D. L., Poppleton, L. E., Ludman, E. J., & Simon, G. E. (2010). Evaluating the effectiveness of cognitive-behavioral teletherapy in depressed adults. *Behavior Therapy*, 41(2), 229-236. <https://doi.org/10.1016/j.beth.2009.03.002>
- Tompkins, C. A., Scott, A. G., & Scharp, V. L. (2008). Research Principles for the Clinician. In R. Chapey (Eds.) *Language intervention strategies in aphasia and related neurogenic communication disorders* (5<sup>th</sup> ed., pp. 179-181). Lippincott Williams & Wilkins.
- van der Meulen, I., van de Sandt-Koenderman, W. M. E., Duivenvoorden, H. J., & Ribbers, G. M. (2010). Measuring verbal and non-verbal communication in aphasia: reliability, validity, and sensitivity to change of the Scenario Test. *International Journal of Language & Communication Disorders*, 45(4), 424-435. <https://doi.org/10.3109/13682820903111952>
- Wallace, S. J., Worrall, L., Rose, T. A., Alyahya, R. S., Babbitt, E., Beeke, S., ... & Dorze, G. L. (2022). Measuring communication as a core outcome in aphasia trials: Results of the ROMA-2 international core outcome set development meeting. *International Journal of Language & Communication Disorders*. <https://doi.org/10.1111/1460-6984.12840>
- 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. <https://doi.org/10.1044/2019.PERSP-19-00146>
- Wellek, S. (2002). Testing statistical hypotheses of equivalence. *Chapman and Hall/CRC*. <https://doi.org/10.1201/9781420035964>
- Wellek, S., & Blettner, M. (2012). On the proper use of the crossover design in clinical trials: part 18 of a series on evaluation of scientific publications. *Deutsches Ärzteblatt International*, 109(15), 276. <https://doi.org/10.3238/arztebl.2012.0276>
- World Health Organization (2021). *How to use the ICF. A Practical Manual for using the International Classification of Functioning, Disability and Health (ICF)*. [https://cdn.who.int/media/docs/defaultsource/classification/icf/drafticfpracticalmanual2.pdf?sfvrsn=8a214b01\\_4](https://cdn.who.int/media/docs/defaultsource/classification/icf/drafticfpracticalmanual2.pdf?sfvrsn=8a214b01_4)