

Critical Review:

The accuracy of administering the BDAE-3, BNT, and SRP using telerehabilitation to patients post-stroke

Nimrit Kaur Grewal

M.Cl.Sc (SLP) Candidate

University of Western Ontario: School of Communication Sciences and Disorders

This critical review examines the accuracy of using telerehabilitation methods to assess patients with acquired language disorders, in comparison to traditional assessments conducted face-to-face. Four studies are outlined in this paper to analyze the accuracy of administering short forms of the *Boston Diagnostic Aphasia Examination (BDAE-3)*, the *Boston Naming Test (BNT-2nd edition)*, and *story retelling procedure (SRP)*. Overall, results indicate remote administration of language assessments yield results comparable to face-to-face methods, increasing the support of its use as an alternative mode of service delivery. However, methodological flaws must be overcome in order to increase the feasibility of using telerehabilitation in language assessment of patients post-stroke. Recommendations for speech-language pathologists and suggestions for future studies are provided.

Introduction

Traditionally, services in speech-language pathology have been conducted with clinicians and clients meeting face-to-face. However, difficulties have appeared in providing equally accessible services to individuals with enduring communication deficits who require long-term management. Individuals who are susceptible to these issues include those with acquired language disorders as a result of stroke or traumatic brain injury (TBI) for example. Speech-language pathologists (SLP's) may have difficulties prioritizing their caseloads and providing equally accessible services to these individuals if client attendance is affected by factors such as physical limitations or if they are located in remote regions (Theodoros, Hill, Russell, Ward, & Wootton, 2008). Additional issues that may interfere with accessibility to services include those associated with cost, weather, as well as family or work schedules (Canadian Association of Speech-Language Pathologists [CASLPA], 2006).

Such issues may be alleviated by adopting *telerehabilitation* as an alternative method of service delivery. Telerehabilitation (a.k.a *telepractice*) is defined as "the application of telecommunications technology to deliver professional services at a distance by linking clinician to client, or clinician to clinician, for assessment, intervention, and/or consultation" (American Speech-Language-Hearing Association [ASHA], n.d.). Innovations in technology, such as the internet and videoconferencing, as well as increased accessibility and use of computers have made assessment via telerehabilitation more feasible (Georgeadis, Brennan, Barker, & Baron, 2004; Hill, Theodoros, Russell, Ward, & Wootton, 2006; Palsbo, 2007). Use of this alternative service delivery in speech language pathology can benefit patients by improving the frequency, accessibility and efficiency of services, as well as reducing the burden and cost of travel.

Telerehabilitation can also aid collaborative work by allowing members access those with more specialized expertise or services (CASLPA, 2006).

In order for telerehabilitation to be adopted as an alternative service modality in speech-language pathology, standardization of assessment protocols must occur. Additionally, high levels of validity and reliability must be demonstrated by remote administration of language assessments, in providing enough information to guide planning for intervention (Theodoros et al., 2008). In other words, "the quality of services delivered through telerehabilitation must be consistent with the quality of services delivered face-to-face" (ASHA, n.d.).

Objectives

The primary objective of this paper is to investigate the accuracy of administering BDAE-3, BNT-2, and SRP using telerehabilitation to patients post-stroke. Another objective is to propose evidence based research and clinical recommendations regarding the use of this modality in language assessment of patients post-stroke.

Methods

Search Strategy

Online databases (Proquest Education, Medline, PubMed, and CINAHL: Cumulative Index to Nursing and Allied Health Literature) were searched using the terms (telerehabilitation) or (videoconferencing) AND (speech language pathology) or (language) or (assessment) or (speech) or (adults).

Selection Criteria

The search was limited to studies that examined telerehabilitation studies dating from 1999 to 2010 involving adults and standardized tests of language assessment.

Data Collection

The results from the literature search generated one nonrandomized clinical trial and three randomized experimental studies with concurrent cohort controls: (1) double cross-over agreement design (2) two randomized controlled trials.

All studies examined at least one of four measures of language assessment: short forms of the BNT, BDAE and selected subtests (specifically, *Conversational and Expository Speech*, *Commands*, and *Auditory Comprehension*), as well as SRP.

Results

Selected subtests of the BDAE-3

Palsbo (2007) examined the equivalency of assessing individual's functional communication post-stroke in remote (T) versus traditional face-to-face (FF) assessment conditions, by using a randomized double-cross-over agreement design. Two pairs of SLP's were designated to a T or FF condition, and administered the following subtests of the BDAE-3 over a 30 minute session: *Conversational & Expository Speech* (i-1), *Commands* (2-c), and *Auditory Comprehension* (2-d). Twenty-four participants, ranging from 25-81 years of age (time post-stroke: 2 months to 15 years) were randomly assigned to a FF or T group at a rehabilitation hospital in Washington, D.C. or Oklahoma City. In the T group, a videoconferencing bandwidth of 346 kbits/sec was used. One SLP administered the assessment from a remote location, while the other SLP sat near the participant in another area. Both SLP's also rated the participant's performance during the assessment. In the FF group, all methods and procedures were the same as the T condition, however one SLP led the assessment while the other SLP recorded responses and did not assist with any aspects of the assessment.

Results from FF and T settings ranged from 92-100% for percentages within 95% agreement and those within a single point of exact agreement. However, exact percentage of agreement differed from the FF setting (50-67%) compared to the T setting (8-25%), which may be attributed to one of the examiners scoring more conservatively than the others.

A thorough description of criteria required to conduct a telerehabilitation equivalence design was outlined, making the study design easier to reduplicate. Test-retest bias was reduced by clinicians simultaneously rating participants performance, as was subjective bias by randomizing participants. Many flaws were evident in the methodologies, which reduce the external reliability of results. Using a sample of convenience raises the potential for recruitment bias, and lack of

participant description make it harder to generalize results to other patients post-stroke. Researchers were not blinded or randomly assigned to each condition, increasing the chances of subjective and examiner bias. Low content validity is evident, as the length of the sessions and failure to fully administer the BDAE-3 in its entirety are not truly reflective of typical language assessments. Finally, a limited description of study design, procedures, videoconferencing equipment protocols, and SLP's telerehabilitation training were evident as well, making it difficult to replicate the study and provide evidence for remote assessment.

Although the results of this study are optimistic, there are significant methodological flaws that must be overcome in order to fully demonstrate the feasibility and equivalency of using telerehabilitation as an alternative modality to assess language skills.

BDAE-3 and short forms of the BNT-2

Theodoros et al. (2008) examined the validity and reliability of assessing language skills by administering the BDAE-3 and short forms of the BNT using an internet-based videoconferencing system. A randomized controlled trial was conducted, in which 32 participants and two SLP's were randomly assigned to a FF or T group. Individuals diagnosed with acquired aphasia as a result of stroke participated, with the exception of one participant with traumatic brain injury (TBI) who presented with a discrete aphasia syndrome. Ages ranged from 21-80 years, with two months-10 years post-onset. Examiners were blind to severity of participant's aphasia prior to assessment and had no prior experience with them. SLP's were randomly assigned to either lead the assessment or rate the participants performance an equal number of times, as well as provide a diagnosis of type of aphasia. Both SLP's underwent training in administering the language assessments and using telerehabilitation system before commencing the study.

In the T group, two computers were equipped with videoconferencing software using an internet bandwidth of 128 kbits/sec. These computers also had high resolution (640 x 480 pixels), high quality video-footage and audio files for data sharing controlled by the lead SLP (e.g., instructional images, video clips, remote camera control, and touch screen capabilities), and finally NetMeeting security guidelines to ensure patient confidentiality. Participant's responses were captured using two cameras mounted on a robotic arm controlled by the lead SLP, which allowed for videoconferencing and store-and-forward capabilities. Participants also wore a headset microphone to record utterances and headphones to hear instructions from the lead SLP. In the remote location, the lead SLP established a

videoconference link, while in another location the silent SLP oriented the participant in front of the computer and assisted with the headset microphone. The lead SLP administered the assessment by instructing & displaying materials on the participants monitor, using pre-recorded instructions and demonstrations in order to overcome audio or visual break-up. The lead SLP and silent SLP simultaneously scored the participant's responses, either online or at a later time by referring to recorded video and audio files. At the end of testing, participants completed an 8-item satisfaction questionnaire (5-point scale) with assistance from the silent SLP if they had difficulties reading. In the FF group, assessments were conducted in a quiet room using standard test instructions from the BDAE-3. Materials used included a stopwatch for timed tasks as well as a desk microphone, minidisk recorder, and digital video to capture audio and video footage of the assessment. One SLP led the assessment, while the other silently monitored audio and video recordings. SLP's scored either online or at a later time as well.

No significant differences in test scores were found between FF and T groups using Wilcoxon signed ranks tests of difference ($p \leq 0.01$). Quadratic weight kappa coefficients indicated moderate-very good agreement between four assessors in language assessments ($k = 0.59-1.00$), with 75% of subtests and rating scales displaying very good agreement. Between assessors, six types of aphasia were identified (90.6% exact agreement), as well as four levels of aphasia severity (very good agreement). In the T group, good to very good inter- and intrarater reliability was found across the majority of measures, as was also evident in test-retest reliability. However, poor inter-rater reliability was obtained when rating the frequency of paraphasias in speech ($ICC = 0.34$). Descriptive analysis of questionnaires indicated participants had high satisfaction with the T setting, with 93% comfortable with it, 100% were eager to participate again, and 80% would be satisfied if assessments were conducted in either modality. However, only 60% of participants found T conditions more convenient than FF, which may have been due to their lack of computer experience.

Intermittent reduction of audio and visual quality was problematic at times, which could have been due to the internet bandwidth used. In addition, a need to modify questionnaires to suite the comprehension needs of language-impaired individuals was identified, as well as when in the rehabilitation stage telerehabilitation consults would be most appropriate. A thorough description of the methods and procedures was outlined by the authors, making the study easier to reduplicate. Subjective and examiner bias were reduced by randomizing SLP's and participants, as well as blinding

SLP's to type of aphasia. High levels of inter- and intrarater reliability were also evident, adding to the strength of their findings. However, the accuracy of questionnaires is problematic, as responses may have been influenced by the presence of the SLP if assistance was required filling out the form, or the content may not have fully captured patient's perceptions of both assessment conditions. Finally, the limited sample size may have affected the strength of the outcomes.

Overall, the results reported by Theodoros et al. (2008) support the validity and reliability of delivering standardized language assessments of aphasia remotely, and provide further support for using telerehabilitation in speech language pathology. In addition, this study outlines potential protocols for conducting remote assessment in future research.

In a follow-up study, the same authors, Hill et al. (2009) investigated whether participant's severity of aphasia influenced the ability to administer standardized language assessments using telerehabilitation methods. The same set of participants and procedures for conducting assessment in FF and T groups were used, however data was analyzed using alternative methods of analysis. Data was analyzed by grouping participants into three levels of severity: *mild*, *moderate*, and *severe*. In T settings, the severity of aphasia did not impact the ability to accurately assess participant's abilities, although it may have increased challenges in conducting the assessment. Clinical agreement was above 90% for all rating scales, excluding the melodic line in the mild severity group (80%). *Kruskal-Wallis* test in T and FF settings revealed no significant differences between scores on six clusters of the BDAE-3 due to severity. However, a significant difference for severity of impairment was found in naming ($p < 0.02$) and paraphasia clusters ($p < 0.01$). Post-hoc analysis of naming clusters and paraphasia clusters revealed *very good* agreement between both settings in each severity level, with all Kappas greater than 0.81, as 37.5% of participants presented with *moderate-severe* and *severe* levels of impairment. Participant satisfaction was not influenced by the severity of aphasia, as participants in the T setting rated the audio and visual quality as *adequate* to *excellent*, and *satisfied* to *very satisfied* with the T session.

Administration of conversational related subtests and the BNT were more laborious due to sporadic audio break up, as well as when assessing patients with severe aphasia. The limited number of participants and exclusion of participants with profound or global aphasia resulted in a relatively small sample within each severity level, lowering the ability to generalize results. Finally, a major limitation identified was the absence of

a clinician satisfaction survey comparing assessments conducted in each modality. However, despite these limitations the strength of results are enhanced by the thorough description of methods, procedures, and videoconferencing protocol making the study easier to reduplicate. Randomizing and blinding SLP's reduced the potential for subjective and examiner bias, as did test-retest effects by simultaneous scoring of responses. Finally, high levels of inter- and intrarater reliability and internal validity add to the strength of results obtained.

Overall, results of the Hill et al. (2009) study indicate that the reliability and validity of using telerehabilitation to conduct language assessments is not significantly influenced by the severity of language impairment. Additionally, severity showed no negative affects on participant satisfaction with telerehabilitation, with the exception of assessing naming and paraphasia.

Story-Retell Procedure (SRP)

Georgeadis et al. (2004) investigated whether story retelling by adults with acquired brain injuries was affected by assessments conducted in FF or T settings, and whether their feedback regarding assessment conditions differed as well. A non-randomized case-controlled clinical trial was conducted in which two randomly selected story sets, each consisting of three stories, were taken from the *SRP* and administered to 40 participants. Participants consisted of 12 adults with TBI, 14 adults with left cerebrovascular attack (LCVA), and 14 adults with a right cerebrovascular attack (RCVA), ages ranging from 18-70 years, with less than 14 months post onset. Participants were randomly tested in FF and T settings with one of two different story sets.

Participants in the T setting were seated in front of a computer in one room, with the clinician seated in another room. The videoconferencing system used an internet bandwidth of 10 Mbps/sec, displaying scanned drawings on the participants monitor alongside a video window displaying the remote SLP (~2 ½ x 2 inches). Story retellings were digitally recorded and saved as .wav files in PCM format (11.025 kHz, 8 bit, Mono) at 10 kb/sec. At the end of testing, the SLP interviewed each participant as part of the *Participant Exit Survey*. In the FF setting both the participant and SLP were together in the testing room. As stories were played, the SLP placed drawings on a bookstand in front of the participant. Participant's story retells were digitally recorded using the same procedures as the T setting.

After data was grouped according to etiology, a two-tailed paired samples *t*-test showed no significant differences between participants performance, with a high correlation found in their performance in each group ($p = 0.495$; $r = 0.93$). Overall, both CVA groups

performed the same or better in the T setting relative to FF condition (LCVA $p = .991$; RCVA $p = .064$), whereas the TBI group performed more poorly in the T setting ($p = .213$). Results from one-way ANOVA indicate a trend towards significant differences between both settings (Δ_{T-FF} , $p = 0.069$). Likewise, post-hoc analysis also indicated a similar trend between RCVA and TBI groups (mean difference = 4.77, significance = .079). Recorded responses from both settings were individually played and scored by the SLP, by averaging %IU in each setting. High inter-rater agreement of 92.8% was achieved between the examiner and a blinded SLP (81.1-100%). Feedback from the Participant Exit Survey indicated the majority of participants reported no difference between settings, with some participants reporting higher preference for the FF setting (17/40), and interest in using videoconferencing again (34/40). The remaining who responded were participants with a TBI, and were more likely to provide negative responses towards videoconferencing ($p < 0.001$), as well as a lack of interest in using it again.

The ability to replicate the study and generalize findings was enhanced by the thorough description of participants and sample size, as well as the quality of measures used for analysis and high levels of inter- and intra-rater agreement. However, the potential for examiner bias was increased by not blinding the SLP administering the *SRP*. Sporadic audio/video disruptions were problematic at times, which could have been due to the bandwidth used. A limited description of the videoconferencing system was included, lowering the ability to replicate the study. Results are limited to patients with higher cognitive abilities, since individuals with lower cognitive abilities, severe dysarthria and aphasia were excluded from the study. Finally, the validity of responses may be skewed as they were collected by the SLP during the exit survey.

Overall, story-retelling performance by adults with acquired brain injuries was not affected by either assessment condition. Although patients with TBI expressed less favorable views of telerehabilitation, the majority of participants expressed a high level of interest in using this modality in the future.

Discussion and Recommendations

A compelling level of evidence was presented to support the accuracy and use of telerehabilitation to administer the BDAE-3, short forms of the BNT, and *SRP* to patients post-stroke. Despite promising results, additional investigation is warranted, as the application of remote service delivery in speech language pathology is relatively novel and has limited efficacy research.

First, patient candidacy for telerehabilitation needs to be further explored to determine what adjustments can be made to improve the accuracy of assessment. For instance, the accuracy of results and patient satisfaction in patients with TBI may improve by having a clinician present during assessment to redirect their attention and assist them in staying on task. Alternative methods to increase the accuracy results obtained when administering conversational, naming, and paraphasia subtests to patients with severe aphasia should also be considered. Likewise, altering the administration and format of questionnaires would be beneficial to compensate for any deficits in comprehension patients may have, and may increase accuracy of responses.

Refinements in videoconferencing protocol should be further examined to determine the optimal bandwidth, ways to decrease intermittent disruptions in audio and video footage, and ensuring clinicians are properly trained in the use and troubleshooting of all equipment. For future research, it is critical for authors to include a detailed outline of protocols used to conduct remote assessments, in order to increase the ability to replicate their study and provide evidence for using this method of evaluation. Finally, clinicians must ensure patient confidentiality is maintained in online settings, as was done by Theodoros et al. (2008) and Hill et al. (2009).

All studies highlight the need for future research to include more participants, those with global or profound aphasia, as well as those with severe concomitant motor speech deficits, to increase the power and ability to generalize results. It may be useful to investigate the effects of concomitant cognitive deficits on assessment using telerehabilitation as well (Hill et al., 2009).

To increase the external validity of studies, it would be beneficial for researchers to utilize multiple subject designs whenever possible. It would also be useful for researchers to include a detailed outline of their methods and procedures to increase the ability to replicate their study, and give evidence for or against the remote assessment of language skills of patients post-stroke.

Conclusion and Clinical implications

Overall, there is persuasive evidence to support the use and accuracy of administering the BDAE-3, BNT, and SRP to patients post-stroke using telerehabilitation. Due to the novelty of this modality of evaluation in speech language pathology, clinicians are cautioned to carefully assess patient candidacy before conducting remote testing. Based on the evidence presented, assessment of patients with profound or global aphasia, as well as severe concomitant motor speech or cognitive deficits, should not be conducted remotely until further studies

show it is efficacious to do so. Clinicians should ensure they utilize high bandwidths, are trained to operate videoconferencing equipment, and use online security software to maintain patient confidentiality. Finally, it would be helpful to situate a professional with the patient to operate the equipment, put them at ease, re-direct their attention, as well as note their responses.

References

- American Speech-Language-Hearing Association. (n.d.) *Telepractice for SLPs and Audiologists*. Retrieved November 25th, 2009 from www.asha.org/practice/telepractice/
- Dollaghan, C. (2007). *The handbook of evidence-based practice in communication disorders*. Baltimore: Paul H. Brookes Publishing.
- Georgeadis, A. C., Brennan, D. M., Barker, L. M., & Baron, C. R. (2004). Telerehabilitation and its effect on story retelling by adults with neurogenic communication disorders. *Aphasiology, 18*(5), 639-652.
- Hauber, R. P., Vesmarovich, S., & Dufour, L. (2002). The use of computers and the internet as a source of health information for people with disabilities. *Rehabilitation Nursing, 27* (4), 142-145.
- 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.
- Myers, C., Carey, P., Douglas, A., Kinden, S., Kully, D., Laplante-Levesque, A.,...Svitich, K. (2006). Position paper on the use of telepractice for CASLPA speech language pathologists and audiologists. Retrieved November 26th, 2009 from www.caslpa.ca/PDF/position%20papers/telepractice.pdf
- Palsbo, S. E. (2007). Equivalence of functional communication assessment in speech pathology using videoconferencing. *Journal of Telemedicine and Telecare, 13*(1), 40-43.
- Theodoros, D., Hill, A., Russell, T., Ward, E., & Wootton, R. (2008). Assessing acquired language disorders in adults via the internet. *Journal of Telemedicine & e-Health, 14*(6), 552-559.

