

Critical Review:
Is the assessment of language skills via telerehabilitation comparable to face-to-face assessment in adults with acquired brain injury?

Janet Vu
M.Cl.Sc (SLP) Candidate
University of Western Ontario: School of Communication Sciences and Disorders

This review examined the current literature regarding the equivalence of telerehabilitation methods to face-to-face methods of assessing language skills and disorders in adults with an acquired brain injury. Five studies that compared assessment outcomes for aphasia, functional communication, story-retelling ability, and discourse ability between telerehabilitation and face-to-face settings were selected for review. Overall, findings provided suggestive evidence that telerehabilitation methods of assessing language skills in adults with an acquired brain injury are comparable to face-to-face methods.

Introduction

Following an acquired brain injury (e.g., stroke, traumatic brain injury), many individuals experience communication and language impairments (Mashima & Doarn, 2008). These impairments may include aphasia, functional communication, as well as deficits in story retelling and discourse ability. These impairments prevent the individual from communicating effectively with the people around them. Patients with an acquired brain injury often have accompanying physical impairments and other factors (e.g., lack of transportation, financial constraints) that impact mobility and create challenges in their ability to access care (Hill, Theodoros, Russell, Ward, & Wootton, 2009). Furthermore, many individuals live in rural or remote geographical areas that are underserved by speech-language pathologists (Mashima & Doarn, 2008). For a patient with an acquired brain injury (ABI), the assessment of the patient's communication ability is crucial in order to identify his or her deficits and to determine appropriate intervention and recommendations. Management of a communication disorder is a long-term, ongoing process for affected individuals, and thus the access barriers that present themselves pose a difficult challenge for these individuals (Theodoros, Hill, Russell, Ward, & Wootton, 2008).

Telerehabilitation is a service delivery model that uses technology to deliver rehabilitation services remotely. It enables real-time interaction to occur between clinicians and clients using methods that simulate an in-person experience, such as videoconferencing (Georgeadis, Brennan, Barker, & Baron, 2004). It could also include the use of store-and-forward technology in which data and images are captured first and then transmitted to be viewed and interpreted by a clinician (ASHA, n.d.). Telerehabilitation has already begun to be implemented by speech-language pathologists (SLPs) in Canada and

many other countries around the world (Coleman, Frymark, Franceschini, & Theodoros, 2015). Telerehabilitation could be beneficial for individuals living in remote geographical areas or who experience other distance and mobility issues that prevent them from travelling to receive appropriate care. Not being able to treat patients in a timely matter following an ABI may result in negative long-term outcomes, such as depression and unemployment, which result from their impaired communication abilities. The consequences of these outcomes may ultimately prove to be more costly as well (Turkstra, Quinn-Padron, Johnson, Workinger, & Antoniotti, 2012). Additionally, telerehabilitation could address the issue of generalization of treatment, since methods such as videoconferencing could allow for intervention to occur in the individual's actual life contexts (e.g., at home) rather than the typical treatment setting (i.e., hospital or clinic) (Turkstra et al., 2012).

Therefore, telerehabilitation could potentially serve as a feasible and cost-effective alternative to delivering speech-language pathology services to individuals by overcoming access barriers to care. Specifically, it could be an effective method for conducting language assessments in individuals with communication disorders. However, it is important that assessments conducted via telerehabilitation are equally valid and reliable to assessments performed in-person.

Objectives

The primary objective of this critical review is to evaluate the current literature regarding the equivalence of telerehabilitation methods to face-to-face methods of assessing language skills and disorders in adults with an acquired brain injury. The secondary objective of this paper is to make suggestions regarding clinical implications and future research.

Methods

Search Strategy

Computerized databases, including Scopus, CINAHL, and PubMed were searched using the following terms:

(assessment OR evaluation) AND
(communication OR aphasia OR language)
AND (disorders OR skills) AND
(telerehabilitation OR videoconferencing OR
telepractice)

A systematic review was identified from the above search strategy, and the reference list from this study was hand-searched to identify additional articles for inclusion.

Selection Criteria

Studies selected for inclusion in this review investigated the impact of telerehabilitation on the assessment of language skills in adults with a history of an acquired brain injury. Studies that assessed outcomes of expressive language and receptive language skills were selected. Only studies containing primary research were included and they included assessment measures obtained via both telerehabilitation and face-to-face methods that allowed for comparisons to be made between the two methods.

Data Collection

Five articles that were consistent with the aforementioned selection criteria were selected from the results of the literature search. Two articles described single group studies and three articles employed randomized crossover studies.

Results

Single Group Designs

Single group designs are appropriate methods for comparing the validity and reliability of an assessment method of interest (e.g., assessment conducted via telerehabilitation) to a gold standard method (e.g., face-to-face assessment) (Hill, Theodoros, Russell, Ward, & Wootton, 2009). These methods can be applied to evaluation tools that are used to assess communication disorders. The two following studies employ a cross-sectional, single-group repeated measures design in which novel assessment methods are compared against a consistently applied reference standard, but with the lack of blinding, they are considered to provide level 3 evidence (Howick et al., 2011).

Georgeadis, Brennan, Barker, and Baron (2004) used a single group, repeated measures design to investigate the impact of telerehabilitation on story retelling performance in adults with a history of an acquired

brain injury who were currently receiving SLP therapy. Forty patients were recruited, and the etiology of their injury had been determined to be one of the following: traumatic brain injury (TBI), left-hemisphere cerebrovascular accident (LCVA), or right-hemisphere cerebrovascular accident (RCVA). Although the age range of the sample was large (18 to 70 years), all participants had experienced their injury recently (within 14 months of baseline). Assessments were conducted with each participant in both face-to-face and telerehabilitation settings by the same SLP. All participants were assessed twice (once in each setting). The order of environments in which the participants were tested was randomized to minimize test-order bias; however, the length of time that passed between the two test periods was not specified. Two story sets from the Story Retell Procedure (SRP) were randomly selected and the Percent Information Unit (%IU) was the standardized scoring metric used to assess performance on the SRP. The same SLP who had administered the SRP scored the participants' responses and a second SLP blinded to the testing environment scored a small (5%) sample of these responses; an interrater agreement of 92.8% was reported.

Appropriate two-tailed paired samples t-tests were applied to mean %IU scores across all participants during data analysis and no significant difference was found between performances across the two settings. A significant correlation was also reported between performances in both settings ($r = 0.93$). In addition, data was stratified post-assessment based on the three etiology groups. Results revealed no significant differences within each group; however, the TBI group exhibited poorer performances in the telerehabilitation environment compared to the face-to-face setting. Despite the large age range across the participants, they were similar at baseline regarding the time that had elapsed following their acquired brain injury.

Overall, the study provides suggestive evidence that story retelling performance is similar between telerehabilitation and face-to-face settings. However, the results can only be applied to story retelling tasks in patients without severe aphasia or cognitive-communicative impairment, and this limitation was acknowledged by the authors.

Turkstra, Quinn-Padron, Johnson, Workinger, and Antoniotti (2012) conducted a single group, repeated measures study to determine whether the assessment of discourse ability via telerehabilitation methods is similar to face-to-face assessment. Twenty participants who had a history of traumatic brain injury were recruited, and they ranged in age from 21-69 years of age with time since injury reported as ranging from 1.4 to 29 years.

Tasks from the Mediated Discourse Elicitation Protocol were used in combination with protocol and stimuli from AphasiaBank to obtain assessment measures on four types of discourse: personal event description, story retelling, picture description, and procedural narrative. Each participant was assessed twice: once in each condition by the same SLP, although the time that elapsed in between the two different assessment methods was not specified. The order of setting in which participants were assessed was randomized and the order of stimuli presented was randomized between the two conditions as well in order to minimize systematic bias. A second SLP was responsible for transcribing and analyzing the language samples. A third SLP blinded to the settings in which the participants were tested analyzed a portion of the data (for 20% of participants) to determine a level of percent agreement. Appropriate paired t-tests were performed on the number of C-units, type-token ratio, number of mazes, and number of words used by the clinician in the analysis of the language samples.

Inter-rater reliability for c-unit segmentation was considered adequate (the level of percent agreement between the two SLPs was 92%). However, a Kappa statistic would have provided a more robust measure of agreement, and inter-rater reliability was tested for only a few participants from the already small sample size, a limitation that was acknowledged by the authors. Overall, the results revealed no statistically significant differences in any of the discourse measures across the two settings. Another limitation of this study included the use of non-standardized informal assessment tools for which the reliability and validity have not been widely tested. Due to the selected protocols for this study, it would not be easily replicated. Therefore, this study provides somewhat suggestive evidence that the assessment of discourse skills via telerehabilitation methods is similar to assessment conducted face-to-face. Results of the study can only be applied to participants without severe cognitive impairment.

Randomized Crossover Designs

A randomized crossover design is another appropriate approach to evaluating the validity and reliability of a novel assessment method and comparing it an existing gold standard. The three following studies employ a randomized crossover design in which participants are randomized to groups that are either led by a SLP in the telerehabilitation setting or by a SLP in the face-to-face setting. Participants are simultaneously assessed via both the method of interest (i.e., telerehabilitation assessment) and the method designated as the consistently applied reference standard (i.e., face-to-face assessment). As blinding is not employed in these

studies, they are considered to provide level 3 evidence (Howick et al., 2011).

Palsbo (2007) conducted a randomized, double-crossover agreement study to investigate whether assessment measures of functional communication in telerehabilitation settings were comparable to face-to-face settings. A small number of participants (24) with a history of stroke were recruited through convenience sampling, and there was high variability in time since stroke ranging from two months to 15 years. Three functional communication measures (FCMs) were selected: motor speech, spoken language expression, and spoken language comprehension. The National Outcomes Measurement System (NOMS) was used as the measurement tool to obtain participants' FCMs. Additionally, a small subset of the Boston Diagnostic Aphasia Examination (BDAE) was used to assess speech comprehension, and participants were asked a set of open-ended questions to assess motor speech and spoken language expression. The study took place at two different sites with two SLPs located at each site. At each site, one SLP was designated as the face-to-face assessor and the other as the remote assessor. At each site, participants were randomized into two groups (one group was given the BDAE by the face-to-face SLP and the other by the remote SLP). Participants' FCMs were assessed throughout the entirety of the interaction session simultaneously by the two SLPs they were assigned to (one SLP in either condition).

Percentage agreement calculations were determined and reported, however, a Kappa statistic would have provided a more robust measure of agreement. Other limitations included poorly described methodology that would be difficult to replicate and a high variability in post-onset times across participants. Criteria and rationale were provided for the selection of assessment tools used in the study, however another limitation was the informal use of the BDAE's subtests. Although the BDAE is a widely accepted standardized assessment tool when administered in its entirety, administering only a few subtests does not give a comprehensive depiction of an individual's communication abilities. Results revealed that percentage exact agreement between the SLPs in each setting was lower when assessments were led by the remote SLP (8% to 25%) compared to when administration of the BDAE was conducted by the face-to-face SLP (50% to 67%). The percentage within 95% limits of agreement was much higher, ranging from 92 to 100% when assessments were led in either condition. Furthermore, the two SLPs at each site were not randomized between the two conditions (i.e., one SLP was designated as the face-to-face assessor for all participants), which may have created some bias. However, due to the numerous

limitations, this study provides equivocal evidence for the equivalence of assessing functional communication via telerehabilitation compared to face-to-face assessment.

Theodoros, Hill, Russell, Ward, and Wootton (2008) conducted a randomized, crossover study in which 32 patients with previously diagnosed aphasia were recruited. Participants ranged in age from 21-80 years of age and time since brain injury was reported as ranging from 1 month to 10 years. Both participants and SLPs were randomly assigned to either a telerehabilitation-led or face-to-face-led assessment in order to eliminate test bias. For each participant, a SLP led the assessment in one condition while the SLP in the other condition served as a silent scorer and did not assist in the administration of the assessment. The short forms of the Boston Diagnostic Aphasia Examination, 3rd edition (BDAE-3) and the Boston Naming Test, 2nd edition (BNT), which are both standardized assessments tools for aphasia, were used. Scores on these tests were obtained simultaneously by both SLPs for all participants and then were compared to determine the strength of agreement. The type of aphasia and level of severity were also determined and compared. Inter- and intra-rater reliability of the telerehabilitation SLP-led assessment was evaluated by having two additional SLPs not involved in study to score randomly selected sessions that had been previously recorded.

Overall, good inter-rater and intra-rater reliability were found. Appropriate Wilcoxon signed ranks tests of difference were performed to compare the scores obtained across the two settings, and no significant differences in the BDAE-3 and BNT scores obtained between the telerehabilitation and face-to-face environments were reported. Quadratic weighted kappa coefficients were used to determine the strength of agreement, and results revealed moderate to very good agreement between the two SLPs. Furthermore, a high level of exact agreement was reported between the two SLPs in determining the type of aphasia and level of severity. The limitations in this study included a smaller sample size with participants ranging largely in age and in time post-injury. This study had many strengths including randomization of participants and conditions, well-designed methods, and valid statistic manipulations.

Overall, this study provides suggestive evidence that the assessment of language skills conducted via telerehabilitation is comparable to those conducted in-person. However, these results should only be applied to formal standardized tests of aphasia (specifically the BDAE-3 and BNT) that have been widely used and accepted.

Hill, Theodoros, Russell, Ward, and Wootton (2009) employed a randomized, crossover design to determine whether aphasia severity has an influence on the ability to assess language disorders via telerehabilitation methods. Thirty-two participants ranging from 21 to 80 years of age who had experienced a stroke or traumatic brain injury 2 months to 10 years prior to baseline were recruited. Participants were randomized to either a face-to-face-led or a telerehabilitation-led assessment. For each participant, a SLP was randomly assigned to lead the assessment in one setting while the other SLP became as a silent rater. The short of the BDAE-3 was administered to each participant. This tool specifically consisted of 25 subtests, the short form of the BNT, and seven rating scales. During the assessment, the two SLPs simultaneously scored the BDAE-3, and based on the scores, the SLPs assigned the participant a severity level. Based on the assessment results, participants were grouped into mild, moderate, and severe levels of aphasia. Percentage levels of exact agreement and percentage levels of clinical agreement (i.e., within one scale point) were determined within each severity level during data analysis. Results from the BDAE-3 subtests were grouped into eight language clusters and the Kruskal-Wallis test was applied to the differences between scores obtained in the face-to-face condition and telerehabilitation condition to determine whether the differences were significant. A weighted quadratic Kappa was applied within each severity level if it was found that severity of language impairment had a significant effect.

Results revealed that although percentage levels of exact agreement on the rating scales ranged from 30% to 93.3% percent, the percentage levels of clinical agreement on the majority of speech characteristics of the rating scales were above 90%. No significant differences were reported for the eight language clusters except for the naming and paraphasia clusters. Quadratic weighted Kappa analysis of these two clusters revealed good agreement between the two methods of assessment within each severity level. Strengths of this study included use of the widely accepted standardized assessment tool (the BDAE-3), appropriate randomization of assessor and assessment condition, and the valid statistical manipulations that were applied to the data. A strength of this study was its exploration of the impact that different severity levels of aphasia may have on telerehabilitation assessment. The authors also acknowledged limitations of the study, including the small sample size for each severity level.

Overall, this study provides suggestive evidence that the assessment of language skills using a standardized tool via telerehabilitation is comparable to face-to-face

assessment. The study suggests that although severity of aphasia may have an effect on the ability to accurately assess certain language clusters (i.e., naming and paraphasia), a good strength of agreement is still found within each severity level between the two assessment methods.

Discussion

Overall, findings provided suggestive evidence that telerehabilitation approaches are comparable to face-to-face approaches when assessing language skills in adults with acquired brain disorders. Five articles were selected for review, two of which conducted single group, repeated measures studies. In these studies, participants were assessed twice, once in each type of setting (telerehabilitation or face-to-face). The three other studies conducted randomized crossover studies, in which participants were randomized to a telerehabilitation-led or a face-to-face-led assessment, and they were scored simultaneously by an SLP in each setting. These study designs were appropriate for evaluating the equivalence of telerehabilitation to face-to-face methods of assessment. Considering the methodology employed by the five studies, they provided level 3 evidence.

One of the limitations for the selected studies was the small sample size of participants and the use of convenience sampling in the recruitment process. Due to these limitations, findings from these studies may not be representative of the entire target population. A limitation that was predominant in the studies was a participant pool that was not similar at baseline. Most of the studies had a sample of participants who ranged widely in age. Participants also varied significantly in terms of the amount of time that had elapsed since their injury. Results from the study by Georgeadis et al. (2004) revealed that patients with TBI were more likely to exhibit poorer performances in the telerehabilitation setting. The authors suggested that patients with TBI often have attention deficits that make the assessment process more difficult, and thus with additional materials, technology, and equipment introduced in the telerehabilitation setting, the patients may be in a more challenging environment. These findings suggest that various patient characteristics could have an impact on the ability to assess language skills via telerehabilitation, and therefore the influence of these different characteristics are important to explore. Another recurring limitation for the selected studies was the lack of blinding employed. Neither clinicians nor participants were blinded to the assessment condition; however, due to the nature of the methodology employed, it is a limitation that is difficult, if not impossible, to avoid.

Despite its few limitations, such as the small number of participants recruited, the study conducted by Hill et al. (2009) provided the strongest evidence supporting the use of telerehabilitation. Appropriate and well-designed procedures and materials were used, randomization of assessment setting was employed for both clinicians and participants, and appropriate statistical manipulations were applied to the data. The Boston Diagnostic Aphasia Examination, 3rd edition (BDAE-3) was the widely accepted and standardized assessment tool used, and thus it provided a more robust method of evaluation than the informal assessment protocols used in the other studies. Finally, this study was the only one in which data was analyzed for each severity level in order to determine the impact of aphasia severity on the ability to assess language skills via telerehabilitation.

Limited research has been conducted on the efficacy of telerehabilitation when assessing language skills in adults. Future research could further explore the use of telerehabilitation in the assessment of patients with different severity levels of language impairment as well as the assessment of other communication skills (e.g., voice disorders, motor speech skills). Future studies should make efforts at recruiting larger samples to improve representation of the population of interest as well as grouping participants according to certain population characteristics (e.g., age group, etiology of injury, severity of injury and impairment, etc) to control for confounding factors.

Clinical Implications

Considering the suggestive evidence that assessment results obtained via telerehabilitation methods are equivalent to those acquired in-person, telerehabilitation may be a very viable alternative when face-to-face options of service delivery have been explored. Speech-language pathologists can implement telerehabilitation in their practice in order to provide services to individuals who live in remote areas or face challenges with transportation to communication service providers. Specifically, there is evidence supporting the use of telerehabilitation in the assessment of language functions in adults with an acquired brain injury. However, there are considerations that must be taken into account when implementing telerehabilitation.

Although the use of telerehabilitation has the potential to be equally effective as face-to-face interactions, patients' perspectives are of the utmost importance in the decision-making process. Three of the studies in this review (Georgeadis et al., 2004; Hill et al., 2009; Theodoros et al., 2008) included a participant satisfaction questionnaire, in which participants rated

and commented on their experience with telerehabilitation following their assessment. Overall, participants reported a high level of satisfaction with telerehabilitation methods of assessment and expressed their willingness to participate in telerehabilitation services again in the future. However, in the study by Georgeadis et al. (2004), participants with TBI were less likely to provide positive feedback about their experience with telerehabilitation. The authors suggested that participants' awareness of their poorer performance in the telerehabilitation setting may have contributed to negative attitudes towards telerehabilitation. Participants who had experienced a stroke also provided mixed feedback; although some participants felt more comfortable and felt that assessment tasks were easier in the face-to-face setting, others preferred telerehabilitation. Therefore, participant attitudes towards telerehabilitation are a significant factor when considering the use of telerehabilitation to provide services (Georgeadis et al, 2004).

Hill et al. (2004) discussed the challenges that SLPs may be faced with when using telerehabilitation. The clinicians involved in the study expressed that administration of the assessment tests via telerehabilitation was often more difficult. For example, they had to make certain modifications to the administration protocols, and scoring was challenging because the clinicians had difficulty determining whether the participants were presenting with reduced receptive language skills or whether they were unable to hear the test questions correctly due to audio and video break-up. Furthermore, higher severity of aphasia may increase the challenges with assessment administration

via telerehabilitation, despite similar assessment outcomes (Hill et al., 2004). Formal, standardized tests appear to be the most appropriate to administer via telerehabilitation, as they have already been tested for validity and reliability in addition to lending themselves well to being converted to an electronic format, and the current evidence supports their use. For example, materials from the BDAE can be converted to a computerized version in which participants can use a touch screen to select their responses (Theodoros et al., 2008). The quality of equipment and connection speed are additional factors to consider when implementing telerehabilitation services, as efforts should be made to acquire high-quality video and audio equipment and other necessary resources in order to reduce the possibility of connection break-up. The inclusion of adequately trained staff is also essential, as additional personnel are required to set up the testing environment appropriately as well as handle test materials correctly in-person, while the clinician administers the test via telerehabilitation. Telerehabilitation services should undergo ongoing evaluation in order to continually assess its efficacy.

In conclusion, although there is evidence supporting the use of telerehabilitation in delivery of communication services, more research on its efficacy is needed, particularly when it is used with different population characteristics (such as age, etiology of brain injury, time elapsed since injury, etc). as well as with specific types of assessment (informal versus formal assessment, assessment of motor speech skills versus language skills, etc.).

References

- American Speech-Language-Hearing Association [ASHA]. (n.d.). *Telepractice*. Retrieved from <http://www.asha.org/Practice-Portal/Professional-Issues/Telepractice/>
- Coleman, J., Frymark, T., Franceschini, N., & Theodoros, D. (2015). Assessment and treatment of cognition and communication skills in adults with acquired brain injury via telepractice: a systematic review. *American Journal of Speech-Language Pathology*, 24, 295-315.
- 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.
- Hill, A., Theodoros, D., Russell, T., Ward, E., & Wootton, R. (2009). The effects of aphasia severity on the ability to assess language disorders via telerehabilitation. *Aphasiology*, 23(5), 627-642.
- Howick, J., Chalmers, I., Glasziou, P., Greenhalgh, T., Heneghan, C., Liberati, A.,...Hodgkinson, M. (2011). The Oxford 2011 levels of evidence. *Oxford Centre for Evidence-Based Medicine*. Retrieved from <http://www.cebm.net/index.aspx?o=5653>
- Mashima, P. & Doarn, C. (2008). Overview of telehealth activities in speech-language pathology. *Telemedicine and e-Health*, 14(10), 1101-1117.
- Palsbo, S. (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. *Telemedicine and e-Health*, 14(6), 552-559.

Turkstra, L., Quinn-Padron, M., Johnson, J., Workinger, M., & Antoniotti, N. (2012). In-person versus telehealth assessment of discourse ability in adults with traumatic brain injury. *The Journal of Head Trauma Rehabilitation*, 27(6), 424-432.