



March 2023

Comparing Two Methods of Clinical Supervision

Heather K. Anderson

Louisiana State University, hande1@lsuhsc.edu

Sandra Hayes

Louisiana State University, sandra.hayes@lsuhs.edu

Julie P. Smith

Louisiana State University, julie.smith@lsuhs.edu

Stephanie Currie

Louisiana State University, stephanie.currie@lsuhs.edu

Follow this and additional works at: <https://nsuworks.nova.edu/ijahsp>



Part of the [Communication Sciences and Disorders Commons](#), and the [Other Educational Administration and Supervision Commons](#)

Recommended Citation

Anderson HK, Hayes S, Smith JP, Currie S. Comparing Two Methods of Clinical Supervision. *The Internet Journal of Allied Health Sciences and Practice*. 2023 Mar 20;21(2), Article 10.

This Manuscript is brought to you for free and open access by the College of Health Care Sciences at NSUWorks. It has been accepted for inclusion in *Internet Journal of Allied Health Sciences and Practice* by an authorized editor of NSUWorks. For more information, please contact nsuworks@nova.edu.

Comparing Two Methods of Clinical Supervision

Abstract

Purpose: The purpose of this study was to compare two methods of clinical supervision, remote and in-person, in group speech-language therapy in a university-based outpatient clinic, in terms of student clinical behaviors that occur with each method. Clinical supervision is an essential part of graduate education programs in speech-language pathology and clinical supervisors in university-based speech and hearing clinics often employ remote supervision. The professional literature concerning the impact of remote supervision on student clinical behaviors is limited. The studies found advantages and disadvantages to in-person and remote supervision and no real preference for one method was noted by participants.

Method: This study compared student clinical behaviors, specifically percentage of interactional time, in remote versus in-person supervision conditions. **Results:** No significant difference was found between the two conditions. **Conclusions:** Both in-person and remote are valuable methods for clinical supervision of graduate students. There are strengths and weaknesses to both methods and graduate clinicians tend to present similar behaviors when being supervised via either method.



The Internet Journal of Allied Health Sciences and Practice

Dedicated to allied health professional practice and education

Vol. 21 No. 2 ISSN 1540-580X

Comparing Two Methods of Clinical Supervision

Heather K. Anderson

Sandra Hayes

Julie P. Smith

Stephanie Currie

Louisiana State University Health Sciences Center Shreveport

United States

ABSTRACT

Purpose: The purpose of this study was to compare two methods of clinical supervision, remote and in-person, in group speech-language therapy in a university-based outpatient clinic, in terms of student clinical behaviors that occur with each method. Clinical supervision is an essential part of graduate education programs in speech-language pathology and clinical supervisors in university-based speech and hearing clinics often employ remote supervision. The professional literature concerning the impact of remote supervision on student clinical behaviors is limited. The studies found advantages and disadvantages to in-person and remote supervision and no real preference for one method was noted by participants. **Method:** This study compared student clinical behaviors, specifically percentage of interactional time, in remote versus in-person supervision conditions. **Results:** No significant difference was found between the two conditions. **Conclusions:** Both in-person and remote are valuable methods for clinical supervision of graduate students. There are strengths and weaknesses to both methods and graduate clinicians tend to present similar behaviors when being supervised via either method.

Keywords: clinical supervision, supervision, telesupervision, speech-language pathology

INTRODUCTION AND BACKGROUND

As in other allied health professions, student training in speech-language pathology (SLP) includes both didactic (i.e., classroom) and clinical components. The clinical component is designed to enable students to achieve clinical competencies required for professional practice. To earn a master's degree in SLP, the entry level degree for the profession, students must earn a total of 400 clinical practicum hours and must demonstrate competence in eight practice areas with both pediatric and adult populations. Students in many SLP graduate training programs participate in clinical practicums at university-based outpatient clinics under the guidance of clinical supervisors who provide instruction and training in clinical methods. With increased availability of technology-assisted methods of supervision over the past two decades, many university SLP clinics have incorporated the use of closed-circuit television systems (CCTV) to allow supervisors to directly observe therapy and diagnostic sessions, record sessions, and provide feedback regarding student performance without entering the room. This allows for less distractions for both clients and students and may have other benefits such as fostering increased autonomy and a greater sense of independence in student clinicians.¹ However, the benefits and limitations of this type of supervision compared to having the supervisor in the room with the student and client are not well understood. Given the increased focus on quality of clinical supervision over the past several years, a more robust evidence base is needed concerning optimal supervisory practices.²⁻⁴

The literature concerning use of electronic supervision, particularly in the field of speech-language pathology, is sparse. Carlin et al selected 100 graduate students to survey their perception of in-person versus electronic supervision; of these, 52% completed the survey. The majority reported that e-supervision was convenient and indicated they were supervised more than when supervised in-person. A large percentage (60%) deemed the e-supervision method as a facilitator of their skills and independence; they had to complete all tasks rather than depending on an on-site supervisor. Although the students, overall, did not prefer one method over the other, the students reported more frequent supervision, less stress, and greater convenience with e-supervision.

In another study, Carlin et al investigated the use of e-supervision methods by university liaisons with students placed in offsite practicum in a mixed-methods study involving surveys and interviews with 13 second-year graduate students. Results indicated positive student responses to this supervision. Fewer distractions for clients, increased sense of independence, and scheduling ease were reported as benefits, while lack of personal contact was identified as a limitation.¹

Anecdotally, additional articles did not provide data but provided advantages and suggestions for using e-supervision. Two of these articles' use of e-supervision allowed the supervisor to provide immediate feedback to the student, which they found to be an advantage. Beiting and Nicolet used the chat box feature to provide feedback, whereas Dudding and Justice used an in-the-ear device, allowing students to receive feedback in real time.⁶⁻⁷ In a third article, graduate students stated that although e-supervision provided a feeling of autonomy and flexibility more so than face-to-face, the students suggested maintaining some level of continued face-to-face supervision and interactions. The amount of face-to-face should depend on student experience level and client needs.⁸

While the Carlin et al studies addressed telesupervision in SLP, this study differs from the present study in the method of investigation and the context of supervision.^{1,5} The present study investigated telesupervision with on-site students in the university clinic, with supervisors present in the building, and employed direct observation of behaviors as a data source, in contrast to Carlin's student survey results.

The primary purpose of the present study was to compare two methods of clinical supervision in individual and/or group speech/language therapy in a university-based outpatient clinic in terms of student clinical behaviors that occur with each method. Specifically, the study aimed to characterize student clinical behaviors in two commonly employed scenarios: (a) supervisor present in the same room as the student and patient; and (b) supervisor present in the building but observing via closed-circuit television. The hypothesis was that students would demonstrate a higher percentage of interactional time with clients when a supervisor was physically present in the same room as the students (Condition A) than when the supervisor was remote, defined here as in the same building but supervising from a different room (Condition B). Another goal of the study was to characterize benefits and limitations of a specific quantitative data collection procedure, whole-interval recording, in measuring student clinical performance.

METHODOLOGY

This study was a non-experimental mixed-methods design. The quantitative component was a comparative analysis of student interactional behaviors with pediatric patients under two different conditions. The independent variable was the condition, with or without supervisor physically present in the room, both of which are consistent with common supervisory practices in SLP programs. The qualitative component involved analysis of comments made by the three researchers while observing the participants and recording data on interactional behaviors.

Participants were chosen from graduate students enrolled in their first year of graduate study in speech-language pathology who were assigned to provide group therapy services to two-year-old clients as a course requirement for clinical practicum. The clinical assignment described in this study was a play-based group therapy program for two-year-olds with identified delays in language skills. As per program policy, clinical assignments for these and other students were made based on students' clinical training needs, number of enrolled clients, and student/faculty schedules, with three to five students assigned to this group. All of the student participants started the semester with fewer than 50 clock hours and had no experience outside of the university clinic. The study took place over two semesters. Two students from the group were randomly chosen to be observed for each of the two semesters specifically for this study. All students assigned to the group therapy program were exposed to the same conditions and provided with feedback on their clinical performance in the same manner and at the same frequency.

Prior to the study, the data collection protocol was evaluated for ease of use and reliability. A time-sampling method, interval recording, was selected as the measurement approach because of its usefulness in measuring the presence or absence of selected behaviors over a period of time.⁹ After piloting two interval recording procedures, partial-interval and whole-interval, with various interval lengths, whole-interval recording in 10-second intervals was selected as the most accurate and feasible method. In whole-interval recording, observers record, at the selected interval, whether a specific behavior occurred continuously throughout the interval.

The behavior of interest, clinical interaction, was operationally defined as interaction between the student and one or more clients which is characterized by at least one of the following behaviors: 1) Student spoke to client; 2) Student sat or stood in proximity (within arm's length) to the client and made eye contact with the client or looked directly at the client; 3) Student sat or stood in proximity to the client and looked at objects or actions with which the client was engaged. 4) Student touched the client for the purpose of comforting or redirecting the client (e.g., holds on lap, picks up, pats back, puts arms around client).

Interjudge reliability was evaluated for the whole-interval recording by measuring percentage of agreement between two of the observers in a trial observation session; and the level of agreement was 94%. Prior to the initiation of the study, a fourth researcher, also a faculty member in the SLP program, was added to the team after being oriented to study procedures. This researcher served as a third observer, which allowed for a "majority rule" approach to any disagreements that might occur in the plus-minus ratings: if all three researchers did not assign the same rating, the rating assigned by two of the three researchers was recorded for later data analysis.

Following approval by the Institutional Review Board to carry out the research, data collection was initiated during regularly scheduled group treatment sessions. During all sessions, COVID19 protocols were followed. During both semesters, this included having all clinicians wear masks. During the first (spring) semester, COVID19 protocols also included pairing students with specific children during portions of the group sessions. Throughout the study, the only difference between Condition A and Condition B was the presence of the supervisor in the room. Each student clinician was directly observed for five minutes in each condition by all three observers. The sequence of Condition A and Condition B was alternated from session to session, so that Condition A (supervisor physically present) occurred first during half of the sessions and Condition B (observation via CCTV) occurred first during the other half. To control for factors related to individual supervisor characteristics, all in-person supervision was provided by the same individual, one of the researchers, during the portion of the session being observed for this study.

Observations were conducted during the same portion of the therapy session each time. During each five-minute observation period, behavior was measured by three investigators using the whole-interval recording for time sampling in 10-second intervals, as described above. The Interval Timer IOS application (Apple, Inc.) was used to track intervals by providing a beeping sound every 10 seconds, which gave the observers an auditory cue to record the rating for that interval.¹⁰ Sessions were video recorded, consistent with common supervisory practices in the program, as a backup to live data collection and to allow for reliability checks. Immediately following each observation session, ratings from the three observers were compared. Any disagreements were discussed, and a single composite rating was recorded for each interval. Each single composite rating represented either the rating recorded by all three observers or, as noted above, the rating assigned by two of the three observers. The total number of intervals with a "plus" rating (i.e., interactional time for the whole interval) was recorded for each student for each condition. Comments made by observers during and immediately after data recording sessions were recorded.

To maintain confidentiality, student names were not included on primary data collection forms or files. Students were assigned a number and a list of student participants and corresponding numbers was kept in a password protected file on a password-protected computer within the department. Data used for clinical education (i.e., shared with students as part of standard clinical feedback) were kept confidential in accordance with existing university policies on confidentiality of student grades. Electronic copies of data

records from the study were saved in a password-protected file on a password-protected computer within the department. Paper copies were stored in a locked file cabinet in a limited-access faculty office. Raw data was not transmitted or shared outside of the department. No identifying client information was recorded at any time beyond routine clinical practices employed for all clients.

RESULTS

Data from the whole interval recording consisted of the number of 10-second intervals during which each student was engaged in interactional time for Condition A and Condition B. This yielded 2 data points (interactional time “scores”) per condition per session, for a total of 50 total “scores” per condition (100 scores total). Quantitative analyses were performed to compare mean interactional scores between Condition A and Condition B and to analyze performance of individual student participants. For the group comparison, descriptive statistics included the supervisory score (total number of whole intervals in which interaction was observed for each condition overall, for each condition broken down by academic semester, and for each participant). The overall mean interactional score was 13.86 for Condition A and 12.78 for Condition B, as shown in Table 1. Mean interactional scores by semester are also shown in Table 1.

Table 1. Mean Interactional Scores for Conditions A and B for Each Semester

Time period	Mean Score Condition A	Mean Score Condition B
Spring	14.7058824	12.6470588
Summer	12.0625	13.0625
Overall	13.86	12.78

Table 2. Mean Interactional Scores for Conditions A and B for Each Student

Student	Mean Score Condition A	Mean Score Condition B	Group (Spring or Summer)
1	16.235941	11.0588	Spring
2	13.176	14.235	Spring
3	11.625	13.25	Summer
4	12.5	12.875	Summer

A T-test for dependent means was performed to compare overall mean interactional scores in Condition A to scores in Condition B. The resulting test statistic was 0.839651, which is greater than the value required for rejecting the null hypothesis ($p < .05$), indicating no statistically significant difference between the two groups.

For individual student data, descriptive data included the mean interactional score for each student in each condition, as shown in Table 2. Data were also analyzed by creating line graphs of interactional scores over time for visual inspection of changes in level, trend, and variability, as described by Byiers et al.¹¹ Individual line graphs are shown in Figure 1.

Figure 1. Mean Interactional Time Score.

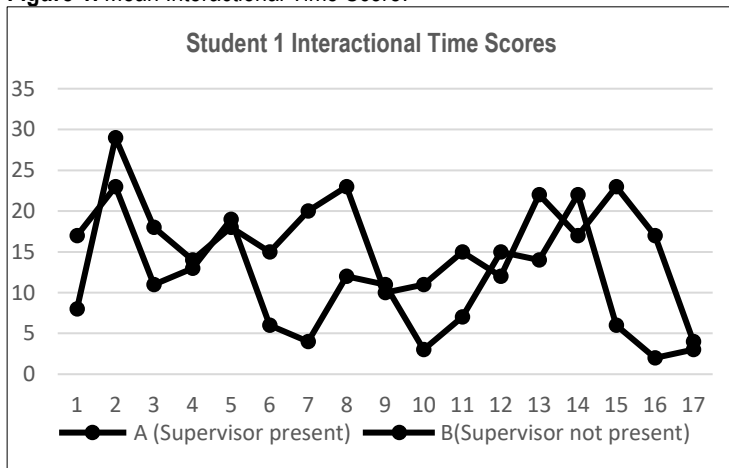
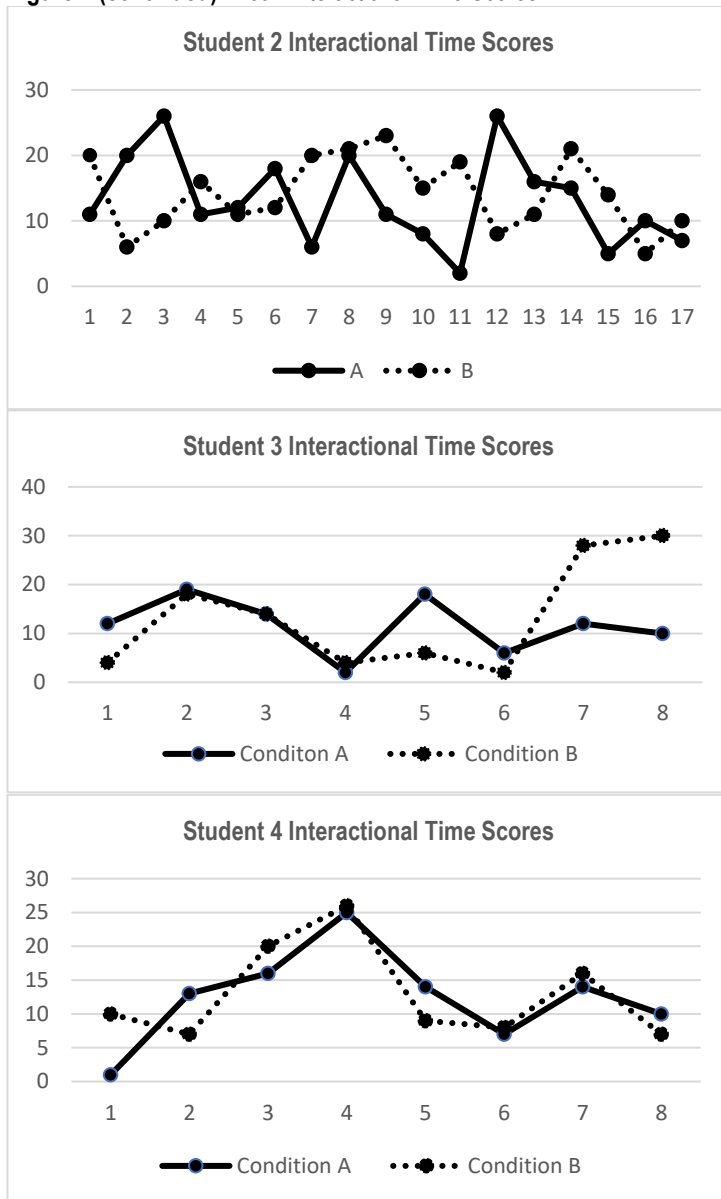


Figure 1 (continued). Mean Interactional Time Scores



Results of the visual inspection of individual student data revealed considerable variability in interactional time scores, particularly in students 1 and 2. As expected with such a high level of variability, the two conditions were not differentiated by upward or downward trend or by level. These observations are congruent with the lack of significant overall difference identified in the combined data, as described previously. To supplement quantitative and visual analyses and allow for evaluation of partial-interval recording as a supervisory tool, qualitative analysis of investigators' comments during data collection was also performed.

Comments were coded to identify major themes using procedures based on those described by Creswell.¹² One theme concerned challenges associated with the logistics of collecting data using the remote supervision system (cameras and closed-circuit TV). For example, investigators noted that the position of students and clients and the level of ambient noise led to difficulty assessing clinical behaviors at times, as evident in comments such as "...could not tell whether she was just sitting there or actually talking," "I wish we could see their mouths better." Another theme was the benefit of the whole-interval recording procedure in helping the observers maintain consistent, careful attention to student behaviors. As one investigator noted, "one difficulty inherent in any 'remote' supervision is sustain[ing] attention, but this [method of data collection] helps." The third and most frequently occurring theme concerned challenges associated with evaluating student behavior relative to the established operational definition of interactional

time. Investigators frequently expressed difficulty classifying behaviors using the binary system required by the whole interval recording due to differences noted even among behaviors that met the operational definition of interactional time. Examples of comments related to this theme include, "I didn't quite know what to do with this one" and "I wasn't sure on some of these." The word "marginal" was frequently included in comments related to this theme, as in, "Meets criteria but marginal" and "the better [the therapy] is, the easier it is [to determine the appropriate rating]."

DISCUSSION

This study aimed to compare the level of student clinical interactions when a supervisor was physically present in the room versus supervision that took place via closed-circuit TV. Quantitative analysis revealed no statistically significant difference in interactional times, as measured by whole-interval recoding, between the two conditions. This finding suggests that the "e-supervision" and in-person models of supervision do not necessarily have a differential impact on student behavior. This is consistent with Carlin et al's finding of a lack of student preference for either condition, and the absence of any detrimental effect of e-supervision found in prior studies.⁵

Key themes identified in the qualitative analysis of supervisor comments included logistical challenges, benefits of whole-interval recording and challenges in evaluating student behavior. Some of the identified logistical challenges were specific to the group clinical scenario. For example, difficulty hearing students' speech due to ambient noise would be less likely to occur in an individual therapy session, which would involve only the client, supervisor and graduate clinician and would likely take place in a smaller room. However, comments related to the impact of student/client positioning on the supervisors' ability to see every aspect of clinical behavior serves as a caution for supervisors employing this method to carefully consider any potential barriers to accurate observation of students, such as sound quality, positioning, and camera angles.

Another key theme identified via qualitative analysis was benefits of whole interval recording. This finding supports the use of the technique for gathering quantitative data concerning student performance. For example, the positive impact of this method on supervisors' sustained attention to the clinical interaction suggests that supervisors may be able to use this tool to improve the accuracy and consistency of their observations of students during clinical training. However, the challenges in evaluating student behavior relative to the operational definition of interactional time reveals the limitations of quantitative measurement, particularly a binary system, in measuring student behavior.

One of the most significant challenges early in the design of this study was construction of an operational definition of interactional time. Three clinical supervisors with considerable experience (i.e., more than 15 years, on average) in clinical supervision of graduate students collaborated on the definition. However, in practice, determining whether a particular interval represented interactional time proved more difficult than anticipated. One possible implication of this difficulty is that the operational definition itself could have been limited. The definition was tested and revised during the early stages of planning, particularly during interjudge reliability testing, and initially appeared satisfactory. Once data collection was in progress, in the interest of internal validity, the operational definition was not modified. It is recommended that in future research of this nature, additional pilot testing of the measurement tool be conducted over multiple sessions.

Another implication is that such difficulty is inherent to evaluation or research involving quantitative measurement of human behavior. While some basic aspects of communication behaviors may be more easily quantified, such as presence or absence of child-directed speech, the nuance and quality of communication cannot be captured in a binary measurement system. The frequent occurrence of the word "marginal" in the comments of the observers (researcher-supervisors) suggests awareness of multiple/varying levels of student performance that could not be easily classified using a binary system. In effect, while the observers complied with the binary classification of behavior required by the quantitative measurement tool, their narrative comments were more consistent with a multilevel classification scheme, in which differences in performance were observed even among students who received the same binary (plus or minus) rating for a given interval. Therefore, the whole-interval recording tool, while beneficial, should not be considered a standalone tool; rather, it is recommended as a supplemental or complementary tool to qualitative description, analogous to the use of both acoustic measures along with auditory-perceptual and visual assessment in the evaluation of voice quality. Another option for facilitating quantitative evaluation that could mitigate the difficulty encountered with this binary rating method is using multi-level (e.g., three-point or five-point) Likert-type scales.

Limitations

Some limitations are noted concerning these findings. One is the limited amount of data for both quantitative and qualitative analysis. Although the number of interactional scores in each condition was sufficient for the statistical analyses, the small number of participants from which the scores were obtained limits generalizability of findings. Likewise, several comments were recorded to

allow for identification of themes in the qualitative analysis additional material, but a broader range of comments or other qualitative data would have allowed for more robust analysis.

Two additional limitations were related to the physical setting and circumstances. First, the clinical procedures observed during this study were subject to the COVID19 protocol of clinicians wearing masks during all sessions. The masks could have impacted the researchers' ability to hear and accurately comprehend the participants' speech. Second, the group therapy took place in a large area compared to standard therapy or exam rooms, which allowed participants to frequently move to different areas of the room. The challenge posed by following this frequent movement with the camera could have impacted the perception and evaluation of visual and auditory information.

Another factor that could have impacted the results was the overall clinical performance of the student participants. Over time, the researchers observed that the students who were randomly selected as participants, using the process described above, demonstrated weaker clinical skills in this context compared to other students with the same level of clinical experience. Although all students were exposed to the same conditions, the fact that target behaviors were measured in weaker students could have impacted the results and may have led to the difficulty encountered by observers in classifying "marginal" clinical behaviors.

CONCLUSIONS

Despite identified limitations, these findings contribute to the body of knowledge concerning effective practice in clinical supervision by providing evidence that student behaviors were not significantly different with or without a clinical supervisor physically present in the room. Additionally, the whole-interval recording method, more typically used in other scenarios such as applied behavior analysis, was found to be a beneficial supplemental tool for quantitative assessment of clinical behaviors. With the increased recognition in recent years of the need to expand the evidence base in clinical education and supervision, additional research is needed to develop and evaluate tools for evaluation of student clinical behaviors. It is recommended that future research focus on evaluation of clinical behavior during individual therapeutic interactions (students interacting with individual patients). Investigation of differences in clinical interactions according to client age or type of communication disorder, such as articulation or fluency, is also recommended. Investigation of the use of other types of quantitative measures in supervision, such as partial interval recording or multi-level rating systems, is also recommended. While individual responses to different supervision models will vary, effective measurement tools, both qualitative and quantitative, are necessary for complete and accurate evaluation of clinical performance.

REFERENCES

1. Carlin CH, Boarman K, Carlin E, Inselmann K. The use of e-supervision to support speech-language pathology graduate students during student teaching practica. *Int J Telerehabil.* 2013;5(2):21-32. <https://doi.org/10.5195/ijt.2013.6128>
2. American Speech-Language-Hearing Association. Knowledge, skills and training considerations for individuals serving as supervisors. [Final report, Ad Hoc Committee on Supervision]. 2013. www.asha.org/siteassets/uploadedfiles/Supervisors-Knowledge-Skills-Report.pdf
3. Council of Academic Programs in Communication Sciences and Disorders. Standards for accreditation of graduate education programs in audiology and speech-language pathology. [White Paper: Preparation of Speech-Language Pathology Clinical Educators]. 2013.
4. Council on Academic Accreditation. 2017. <https://caa.asha.org/wp-content/uploads/Accreditation-Standards-Documentation-Guidance.pdf>
5. Carlin CH, Milam J, Carlin EL, & Owen A. Promising practices in e-supervision: Exploring graduate speech-language pathology interns' perceptions. *Int J Telerehabil.* 2012;4(2):25-38. 10.5195/ijt.2012.6103
6. Beiting M, Nicolet, G. Screenless teletherapy and silent telesupervision: Leveraging technology for innovative service delivery and clinician training in speech-language pathology during the COVID-19 era. *CommonHealth.* 2020;1(3):106-120. 10.15367/ch.v1i3.413
7. Dudding CC, Justice LM. An e-supervision model: Videoconferencing as a clinical training tool. *Commun Disord Q.* 2004;25(3):145-151. <https://journals.sagepub.com/doi/10.1177/15257401040250030501>
8. Dudding C. Distance supervision: An update. *Perspectives of the ASHA Special Interest Groups.* 2006;16:16-17. <https://doi.org/10.1044/aas16.1.16>
9. Cooper JO, Heron TE, Heward, WL. *Applied Behavior Analysis* (2nd Ed.). Pearson Education; 2007.
10. *Interval timer* [Smart Phone Application]. Version 4.4 Cupertino, California. Apple, Inc.; 2021.
11. Byiers B, Reichle J, Symons, F. Single-subject experimental design for evidence-based practice. *Am J Speech Lang Pathol.* 2012;21(4). [https://doi.org/10.1044/1058-0360\(2012/11-0036\)](https://doi.org/10.1044/1058-0360(2012/11-0036))
12. Creswell JW. *Research Design: Quantitative, Qualitative, and Mixed Methods Approaches* (2nd Ed.). Sage Publications; 2003