Analysis of Clinical Supervisor Feedback in Speech-Language Pathology

JoAnne Cascia

Monmouth University
West Long Branch, NJ

Disclosure: JoAnne Cascia has no financial or nonfinancial relationships related to the content of this article.

Abstract

Purpose: The purpose of this study was to measure the interrater agreement of assessment and feedback among speech-language pathology clinical supervisors. Two questions were posed: (a) Does the use of a structured tool that focuses on predetermined target areas of a therapy session increase the interrater agreement of assessment and feedback among a group of clinical supervisors? (b) Do similarities in a clinical supervisor's background and experience result in greater interrater agreement of assessment and feedback of an observed therapy session?

Method: This study used a repeated measures design. Participants included 15 clinical supervisors who observed therapy sessions and collected observational feedback in 2 different ways. Method 1 involved the use of the recording evaluative statements method, and Method 2 involved a structured tool developed by the researcher.

Results: A series of 3 paired-sample t tests were used to compare the means of Method 1 and Method 2, revealing a significant difference between the Methods, with Method 2 resulting in greater interrater agreement.

Conclusion: This study revealed that clinical supervisors are more in agreement with each other regarding the elements of a therapy session when they are given guidelines and a structured way to collect observational feedback. These results suggest the need for training and consistency in regard to assessment and feedback in the area of clinical supervision.

Clinical supervision has remained a critical element in the training of student clinicians for successful careers in the field of speech-language pathology. The feedback provided by supervisors can be instrumental in preparing and teaching students to provide effective therapy. The American Speech-Language-Hearing Association (ASHA, 2008a) position statement on clinical supervision stated, "Clinical supervision is a distinct area of practice in speech-language pathology and it is an essential component in the education of students and the continual professional growth of speech-language pathologists" (Position Statement section).

Clinical supervision in speech-language pathology involves observing a student clinician's therapy sessions, collecting data from the sessions regarding the student clinician's performance, conducting a conference or an informal discussion regarding the observation, and providing performance feedback. The conference between a student clinician and a clinical

supervisor is an opportunity for the student clinician to ask questions, learn techniques, and understand the rationale for specific therapy techniques and activities. This is also an opportunity for the clinical supervisor to provide feedback regarding the student clinician's strengths and weaknesses and to provide suggestions for improving therapy sessions.

According to McCrea and Brasseur (2003),

> Beginning supervisees need to learn that there is a core of categories of clinical behaviors that are important to most clinical interactions. . . . [These core categories can include] recording client behavior, determining reinforcements and reinforcement schedules, giving instructions, establishing and providing cues, providing feedback to clients, staying on task, and other behaviors appropriate to the specific client. (p. 195)

Clinical supervisors should target these specific areas when conducting observations and providing feedback. This study was intended to determine whether the evaluation of predetermined target areas during the clinical observation process would result in greater interrater agreement of performance feedback among a group of clinical supervisors.

Statement of the Problem

At a university in the northeastern United States, speech-language pathology graduate students typically have up to five supervisors from the time they begin their first clinical experience to the time that they complete their clinical fellowship. These may include two supervisors in the on-campus clinic setting, one supervisor for each of two externship settings, and one supervisor for the completion of the clinical fellowship. McCrea and Brasseur (2003) stated, "Before becoming certified professionals, speech-language pathologists or audiologists will have been supervised by many different persons across several kinds of settings" (p. 5). The performance feedback given to a student clinician by his or her supervisor is typically the student clinician's first experience in participating in the supervisor–supervisee dyad in the field of speech-language pathology. The feedback received at these initial conferences creates the foundation on which the student clinician will build a career. In fact, Smith (1984) found that students ranked receiving supervisor observation and feedback as the most effective factor contributing to their skill development. Bischoff, Barton, Thober, and Hawley (2002) found "supervision to have the most enduring influence on the development of clinical self-confidence" (p. 376).

Graduate students in the field of speech-language pathology often received conflicting performance feedback from one supervisor to the next. Inconsistent supervisory feedback can often create confusion for students with regard to appropriate therapy techniques and treatment approaches. Inconsistent and sometimes contradictory supervisory feedback can be detrimental to the learning process of student clinicians as they rely solely on their supervisors for direction. The implications of this situation may be far-reaching and impact the student's remaining clinical placements, clinical fellowship placement, and performance as a licensed and certified speech-language pathologist. London (2003) explained, "When feedback is inconsistent, people's fear of negative information may outweigh their satisfaction with the positive information" (p. 34). Hoben, Varley, and Cox (2007) reported that student speech-language pathology clinicians have difficulty conceptualizing problems, planning diagnostic strategies, organizing information, and interpreting findings. Given that these tasks are difficult for students, it could be assumed that if the performance feedback they receive from their supervisors is inconsistent or contradictory, it may be even more difficult for them to organize, plan, and execute therapy. Furthermore, London (2003) suggested, "to avoid biases that may affect reactions to the feedback and later revision of self-image, it is important that people receive consistent information or at least information that suggests agreement among raters" (pp. 34–35).

At a university in the northeastern United States, supervisors observe sessions and collect observational data using a method that McCrea and Brasseur (2003) refer to as

recording evaluative statements. This method requires supervisors to write statements and take general notes regarding the session they are observing. The notes they write can refer to any aspect of the student clinician's therapy session or performance. Supervisors then conduct a conference with their student clinicians providing feedback about the session that they observed. Student clinicians have occasionally expressed confusion or questioned the feedback they received, as the feedback was, at times, reportedly contradictory to feedback they had received from previous clinical supervisors. In other words, students may be told by their supervisors that a particular aspect of their performance was executed well and should be repeated or maintained; the following semester, a new supervisor then tells them that the same aspect of their performance is not adequate and should be changed. This problem may be occurring because student clinicians are participating in their first clinical experiences and, as stated by McCrea and Brasseur, beginning supervisees need to learn clinical skills in specific "core categories" (p. 195).

Most clinical supervisors, however, are entering the clinic setting with richly diverse experiences, including varied work settings, participation in many continuing education courses, extensive training in specific areas of the field, and distinct personalities. Geller (2002) stated that it is common for supervisors to use their own experiences as a model for their expectations of student clinicians. These are all factors that contribute to the supervisor's assessment and subsequent feedback of a session, sometimes leading each supervisor to a different conclusion.

Pershey and Reese (2002) stated, "Importantly, supervisors and student clinicians should jointly select a small, manageable number of behaviors or conditions to target for change" (p. 202). By utilizing the method of recording evaluative statements, it is difficult for supervisors to target small, manageable, or distinct areas or behaviors for change. In addition, McCrea and Brasseur (2003) cautioned against the habitual use of a method, such as recording evaluative statements, as student clinicians are not provided with a rationale regarding the supervisor's statements. McCrea and Brasseur explained that active, direct behaviors of the supervisor, as evident in the use of recording evaluative statements, "do not encourage, but perhaps, discourage, self-analysis and creative thinking on the part of the supervisee" (p. 167). This is perhaps because recording evaluative statements provides no predetermined target behaviors upon which the clinician can focus and improve. Student clinicians do not know exactly what their supervisors will be evaluating them on until after the evaluation occurs and they are provided feedback. This idea is supported by the statement mentioned earlier by Pershey and Reese (2002), which maintains that supervisors and supervisees should select a "small manageable number of behaviors or conditions to target for change" (p. 202). Using the recording evaluative statements method, supervisors are unlimited as to the number of behaviors or areas that they evaluate when observing a student clinician.

It was important, therefore, to determine whether greater interrater agreement of performance feedback among supervisors could be achieved by recording evaluative statements or by providing clinical supervisors with a structured tool for recording observations that targets predetermined areas of a student clinician's therapy session and provides a rationale for each supervisory comment. According to McCrea and Brasseur (2003), these predetermined areas are based on the core categories of clinical interaction, including "recording client behavior, determining reinforcements and reinforcement schedules, giving instructions, establishing and providing cues, providing feedback to clients, staying on task, and other behaviors appropriate to the specific client" (p. 195) as well as on clinical skills that are measured by other supervisory analysis systems. These systems include the Wisconsin Procedure for Appraisal of Clinical Competence (W-PACC; Shriberg et al., 1974) and the Indiana University Evaluation of Speech-Language Pathology Student Practicum (IUESLPP; Indiana University, 1996). Elements of these evaluation systems are used, rather than either system in its entirety, as each system is inclusive of additional measurements that were not addressed in this study. In addition, because these systems were designed to evaluate student

clinicians' performance over the course of a semester rather than their performance during one therapy session, the systems contained many elements that cannot be measured during the observation of only one session. Target areas from these evaluation systems, in addition to those from McCrea and Brasseur's (2003) core categories, were used to generate the structured evaluation form.

Purpose of the Study and Research Questions

The purpose of the study was to establish whether greater interrater agreement of performance feedback among a group of supervisors can be achieved by (a) recording evaluative statements or (b) providing clinical supervisors with a structured tool for recording observations that targets predetermined areas of a student clinician's therapy session based on the core categories of clinical behavior, and provides a description and rationale for each supervisory comment. To further evaluate interrater agreement, the supervisors' responses were compared on the basis of their background, education, and experience to determine whether one subgroup of supervisors (e.g., full-time faculty members, supervisors with equal years of experience, etc.) had higher interrater agreement among them.

On the basis of the purpose, two research questions guided this study:

1. Does the use of a structured tool that focuses on predetermined target areas of a therapy session increase interrater agreement of assessment and feedback among a group of clinical supervisors?
2. Do similarities in a clinical supervisor's background and experience result in greater interrater agreement of assessment and feedback of an observed therapy session?

Methodology

This study used a repeated measures design (Irwin, Pannbacker, & Lass, 2008) to gather both quantitative and qualitative data regarding interrater agreement of performance feedback among the population of clinical supervisors at one university. Repeated measures design includes one group of participants who were exposed to more than one condition (Irwin et al., 2008).

Participants in this study were recruited from the population of clinic supervisors within a university in the northeastern United States that has a graduate training program in speech-language pathology accredited by the Council of Academic Accreditation (CAA) of the American Speech-Language Hearing Association (ASHA). To be eligible to participate in this study, all participants held a master's or terminal degree in speech-language pathology, held a current Certificate of Clinical Competence in Speech-Language Pathology from ASHA, were employed by the CAA-accredited training program in speech-language pathology, and had at least 1 academic year of supervisory experience in a university setting. Formal course work or training in clinical supervision was not required for participation, nor is it required for employment as a clinical supervisor. The participants' university employment did not have to be devoted exclusively to clinical supervision. Both male and female supervisors were invited to participate. There were a total of 15 participating supervisors.

Each participant completed a Demographic Data Sheet and data collection instruments. The Demographic Data Sheet was designed by the investigator to collect data about chronological ages of the respondents, degrees, employment setting, years of experience as a clinical supervisor, approximate number of hours spent in clinical supervision in an academic term, other university-related responsibilities, membership in professional organizations for supervisors, and formal training in clinical supervision. This information was used for verification of subject eligibility, for descriptive data reporting, and for answering the research questions for this study.

Using written and verbal instruction, each participant completed two data collection instruments in this study. The first was referred to as Method 1. Method 1, recording evaluative statements, utilized blank sheets of paper for the participants to collect data regarding three viewed clinical therapy sessions. This method required supervisors to write statements and take general notes regarding the session they were observing. The notes they wrote could refer to any aspect of the student clinician's therapy session and could include either positive or negative comments.

Method 2 utilized a structured data collection form developed by the researcher that contained quantitative yes and no questions, as well as additional space after each question for qualitative descriptions and rationale to be provided. The questions on the form targeted the core categories of clinical behavior (McCrea & Brasseur, 2003), as well as elements from the W-PACC (Shriberg et al., 1974) and the IUESLPP (Indiana University, 1996).

A set of three clinical training videos were developed for this study. The videos depicted simulated speech therapy sessions conducted by professionals in the field of education who are not speech-language pathologists. The individuals acting as the student clinician on the three videos included a reading specialist, an English as a Second Language teacher, and a preschool special education teacher. These individuals were chosen as they each had a background in education, and they were comfortable interacting with children; none of these individuals had specialized training in speech-language pathology. This scenario was judged to be comparable to first- or second-semester graduate students completing a clinical practicum in speech-language pathology. The sessions included the individual acting as a speech-language pathologist conducting a scripted speech-language therapy session with a child. The adults in the video followed a script developed by the researcher; however, the children in the videos did not follow a script of any kind. The videos included a simulated therapy session for articulation remediation, phonological processing, and language acquisition.

The three clinical training videos represented an articulation session with a 9-year-old client, a phonological processing session with a 7-year-old client, and a language acquisition session with a 2-year-old client.

Procedure

On the day of the study, the participating supervisors were each given two sealed manila envelopes, one labeled Packet 1 and the other labeled Packet 2, with matching precoded numbers written on the outside of each.

Packet 1 contained the cover letter describing the purpose and format of the study; the Demographic Data Sheet; and three sheets of paper labeled Method 1, Session 1; Method 1, Session 2; and Method 1, Session 3. The sheets also contained instructions for the participants. The demographic information sheet and each sheet for Method 1 of the data collection also had the same precoded number that corresponded to the number on the outside of the envelope.

The supervisors were instructed to complete the demographic information sheet first. Next, the researcher explained to the participants that they would be watching three clinical training videos of three speech therapy sessions. The participants were instructed to use the blank sheets of paper, labeled Session 1, Session 2, and Session 3, to take notes just as they would when observing a clinical therapy session in the clinic. These instructions were also printed on the top of the paper.

The participants then viewed the three 10-min clinical training videos. The participating supervisors viewed each session and recorded their observational data on the sheets of paper provided. Supervisors were instructed to take notes regarding the session in the same manner that they do when observing a session in the clinic. No other specific instructions were given in an effort to most closely replicate the current method of collecting data and assessing a

session. This method, recording evaluative statements, is the method that was used at the university's speech and language disorders clinic and was considered Method 1 for this study.

The researcher then explained to the participants that they would watch the same three clinical training videos as they did in Method 1. However, this time they were instructed to use the structured data collection forms. They were instructed to circle either "Yes" or "No" for each question and to provide a rationale for their decision in the space provided after each question.

The same three clinical training videos were used for completion of Method 2. The supervisors were instructed to complete the structured data collection form (Method 2) for each of the 10-min sessions. The structured data collection form contained quantitative yes and no questions, as well as additional space after each question for qualitative descriptions and rationale to be provided. The final question on the form allowed supervisors to state which method they preferred using.

The same clinical training videos were used for Methods 1 and 2 in order for the researcher to make a direct comparison between which method provided the greatest interrater agreement of assessment and feedback among the group of supervisors. Direct comparisons could not be made if different clinical training videos were used for Methods 1 and 2. In addition, all supervisors completed Method 1 first, followed by Method 2. This was done in order to ensure that Method 1 most closely replicated the clinic's current method of assessing a therapy session, collecting data, and providing subsequent feedback without viewing the predetermined target areas of assessment in Method 2. It was the opinion of the researcher that if supervisors were to view the target areas in Method 2 before completing Method 1, the results would be skewed as supervisors may complete the narrative method (Method 1) based on the questions they have just read in Method 2.

Results

It was the investigator's aim to collect data from at least 15 participants at one facility. Invitations to participate were distributed to every supervisor at the facility for a total of 21 invitations. Nineteen supervisors agreed to participate; however, due to scheduling conflicts and other commitments, a total of 15 supervisors participated.

All participating supervisors were actively supervising graduate students who were completing their practicum in the on-campus clinic setting. All participating supervisors were female, although this was not an eligibility criterion. Participants completed a demographic information sheet that included data regarding their education, employment status, and professional experience. Participants included three supervisors who were full-time faculty members, six adjunct faculty members, and six supervisors who did not teach any additional course work (see Table 1). In addition, participating supervisors included 12 individuals whose highest level of education was at the master's level and three individuals whose highest level of education was at the doctoral level (see Table 1).

Table 1. Employment, Education, and Professional Experience of Participating Supervisor in Percentages (N=15)

Category	**Status**
Employment status	
Full time	20
Adjunct faculty/supervisor	40
Supervisor only	40
Education	
Master's of arts/master's of science	80
Doctor of Philosophy/Doctor of Education/ Doctor of Speech-Language Pathology	20
Years of experience as an speech-language pathologist	
1 to 5	6.7
6 to 10	13.3
11 to 15	26.7
16 to 20	13.3
Over 20	40
Years of experience as a supervisor	
1 to 2	26.7
3 to 5	20
6 to 10	26.7
11 to 15	6.7
16 to 20	13.3
More than 20	6.6
Population of experience	
Pediatric	73.3
Geriatric	26.7

Participants also reported having gained experience as a speech-language pathologist in a variety of settings, including seven in public schools, five in early intervention programs, nine in hospitals, three in home care, six in rehabilitation centers, eight in private practice, and four in nursing homes. One supervisor reported having worked in only one setting (public school). The other participating supervisors reported having worked in a variety of settings throughout their careers. Participating supervisors also reported having gained their experience working with a variety of clients, including clients with autism (five), aphasia (five), fluency disorders (two), voice disorders (one), and articulation disorders (seven).

Participants reported varied levels of experience as both a speech-language pathologist and a clinical supervisor (see Table 2). Six participants (40%) had over 20 years of experience as a speech-language pathologist. In addition, participants reported having gained their experience while working with either the pediatric or geriatric populations (see Table 2).

Table 2. Means and Standard Deviations for Method 1 as Compared With Method 2 for Each Video (N=15)

Category	*M*	*SD*	*SEM*
Pair 1			
Video 1, Method 1	2.8667	1.55226	.40079
Video 1, Method 2	8.9333	2.71153	.70011
Pair 1			
Video 2, Method 1	2.6000	1.72378	.44508
Video 2, Method 2	8.5333	2.92445	.75509
Pair 1			
Video 3, Method 1	3.0667	1.75119	.45216
Video 3, Method 2	9.5333	2.99682	.77378

Results of Research Question 1

Does the use of a structured tool that focuses on predetermined target areas of a therapy session increase interrater agreement of assessment and feedback among a group of clinical supervisors? A discussion of the data analysis and findings for Research Question 1 follows.

Data for Method 1 were analyzed by first reading each comment and then determining whether it was a positive comment or a negative comment. Next, the investigator categorized each comment based on the corresponding item from the structured data collection form (Method 2). One example follows: If the participant wrote, "Mr. Potato Head was a great choice to use to keep the child interested," then that comment would be deemed positive and assigned to Category 2, material selection, as the comment corresponds to Item 2 on the structured data collection form, "Were the materials appropriate for the client?" Because Method 1 was open ended, the use of this data collection instrument allowed for comments in an unlimited number of categories. If participants wrote comments, either positive or negative, for categories not included in Table 1, these responses were coded and analyzed using qualitative methods. After each comment was reviewed and categorized by the investigator, the total number of positive comments for each session was counted, resulting in a score of 0–14. If one participant wrote five positive comments that were counted and categorized, they received a score of 5 for that session. This procedure was repeated three times for each participant, once for each session that was viewed.

Method 2, the structured data collection form, required participants to circle "Yes" or "No" for each of the 14 questions pertaining to the session that was viewed. Each question represented one category. Analysis of Method 2 included counting the total number of yes responses. If the participant responded, "Yes," to the question, "Were the materials appropriate for the client?" this response represented positive feedback regarding material selection for that session, and 1 point was assigned. This scoring system was repeated for each of the 14 questions on the form, resulting in a total score of 0–14. If one participant circled, "Yes," five times for one session, the clinician received a score of 5 for that session. This procedure was repeated three times for each participant, once for each session that was viewed.

Findings. The Statistical Package for Social Sciences (SPSS: Student Version 17.0) computer program was used for data analysis of Research Question 1. Data were entered, and the SPSS program was used to compute parametric tests, including a series of analysis of variance tests (ANOVAs) and a series of paired-sample *t* tests.

In order to answer the first research question, a 2 × 3 repeated measures ANOVA (Huck, 2008), using the three videos as the factor and the two methods as the between-

subjects factor, was conducted. The repeated measures ANOVA was used to determine whether there was a main effect for the three videos, a significant difference between subjects, and any interaction between the videos and the subjects. In addition, a series of paired-sample *t* tests were conducted in order to further investigate whether there was significance between methods.

A repeated measures ANOVA revealed no main effect for the factor (the videos watched), $F(2, 56) = 1.386$, $p = .259$. These results indicated that there was no significant difference between the three videos used in the study. In other words, the results obtained from one video were not significantly different than the results obtained from either of the other two videos. Next, a repeated measures ANOVA was conducted and revealed no interaction between the factor (the videos) and the method used, $F(2, 56) = 0.198$, $p = .821$. These results were not significant, indicating that there was no interaction between the factor (the videos watched) and the method used. A repeated measures ANOVA revealed the results between subjects was statistically significant, $F(1, 28) = 78.72$, $p < .01$, with Method 2 outperforming Method 1. These results indicated that the group of supervisors provided more similar comments using Method 2 than Method 1. These results were further explored using a series of three paired-sample *t* tests.

A paired-sample *t* test was used to compare the means of two variables: In this case, the two variables are Methods 1 and 2. The paired-sample *t* test then computed the difference between the two variables and tests to see whether the average difference was significantly different from zero (Huck, 2008). A series of three paired-sample *t* tests were performed in order to further investigate the significance of one method over the other for each video viewed. The first *t* test revealed a significant difference between Method 1 ($M = 2.87$, $SD = 1.55$) and Method 2 ($M = 8.93$, $SD = 2.71$) in Video 1, $t(14) = -10.59$, $p < .01$, with Method 2 outperforming Method 1. Video 2 also revealed a significant difference between Method 1 ($M = 2.60$, $SD = 1.72$) and Method 2 ($M = 8.53$, $SD = 2.92$), $t(14) = -9.68$, $p < .01$, with Method 2 outperforming Method 1. The final pair also revealed a significant difference between Method 1 ($M = 3.06$, $SD = 1.75$) and Method 2 ($M = 9.53$, $SD = 2.99$), $t(14) = -9.48$, $p < .01$, with Method 2 outperforming Method 1. These results revealed that Method 2 resulted in greater interrater agreement among the group of supervisors than did Method 1 (see Tables 2 and 3).

Table 3. Results of the Paired-Sample t Tests Comparing Methods 1 and 2 for Each Video

	Pair 1	**Pair 2**	**Pair 3**
Category	V1 M1 and V1 M2	V2 M1 and V1 M2	V3 M1 and V1 M2
M	−6.06667	−5.93333	−6.46667
SD	2.21897	2.37447	2.64215
SEM	.57293	.61308	.68220
95% confidence interval of the difference			
Lower	−7.29549	−7.24827	−7.92984
Upper	−4.83784	−4.61840	−5.00349
t	−10.58900	−9.67800	−9.47900

Note. V=video, M=method, df=14, p=.000.

These results were further explored by determining percentages of interrater agreement for the number of positive and negative comments generated for each video using both methods.

With respect to Video 1, 86% of supervisors reported that the activity selection was appropriate using Method 2, whereas only 13% reported the same using Method 1. Similarly,

6% of supervisors reported that the clinician remained on task when using Method 1, whereas 86% reported task maintenance when using Method 2. This pattern was repeated for positive comments in all categories and for negative comments in seven of the 14 categories, including activity selection, modeling, utilizing teaching moments, adapting the session as needed, providing feedback, cuing, and homework. With respect to Video 2, 26% of supervisors reported that activity selection was appropriate when using Method 1, whereas 100% of supervisors reported that activity selection was appropriate using Method 2. This was also true for material selection in Video 2, as 53% of supervisors felt it was appropriate when using Method 1 but 100% of supervisors felt it was appropriate when using Method 2. With respect to Video 2, it was evidenced that there was a greater percentage of interrater agreement among the supervisor's positive comments in 13 of the 14 categories and among their negative comments in 12 of the 14 categories when using Method 2. With regard to Video 3, similar findings were revealed.

It is evident, based on the data presented in Table 4, that Method 2 resulted in greater interrater agreement of positive comments in all 14 categories and in negative comments in eight of the 14 categories. These percentages further support the findings from the *t* tests that were reported earlier. On the basis of these findings, the null hypothesis was rejected.

Table 4. Percentages of Supervisory Comments Using Methods 1 and 2

Comment Category/ Item on Method 2	**Video 1**				**Video 2**				**Video 3**			
	Positive comments		**Negative comments**		**Positive comments**		**Negative comments**		**Positive comments**		**Negative comments**	
	M1	**M2**	**M1**	**M2**	**M1**	**M2**	**M1**	**M2**	**M1**	**M2**	**M1**	**M2**
1. Activity selection was appropriate	13	86	6	26	26	100	26	13	26	26	26	26
2. Material selection was appropriate	20	86	46	26	53	100	0	6	60	40	40	40
3. Clear instructions were provided	73	86	33	26	0	46	6	26	26	40	40	40
4. Modeling was provided	13	80	6	40	26	66	46	53	40	20	20	33
5. Teaching moments were utilized	6	40	20	60	40	46	33	66	20	26	26	26
6. Body language/ enthusiasm	73	93	20	20	53	93	0	6	33	13	13	33
7. Session was adapted to client's needs	0	40	0	40	0	73	0	20	0	0	0	20
8. Reinforcements were appropriate	40	93	26	26	6	46	20	53	13	6	6	46
9. Feedback was appropriate	26	66	33	53	6	40	6	60	46	13	13	26
10. Data collection was unobtrusive	0	20	26	13	0	13	13	13	0	13	13	20
11. Behavior management was appropriate	0	33	0	0	0	73	0	6	0	0	0	0
12. Session remained on task	6	86	13	0	6	73	13	26	13	0	0	0
13. Cuing was appropriate	6	73	13	33	26	66	6	33	20	0	0	26
14. Assigned homework was appropriate	0	6	0	20	0	0	0	20	0	0	0	6

Note. M=method.

The final question on the structured data collection form asked supervisors to reveal which method of data collection they preferred. Sixty percent of supervisors stated that they preferred the structured form. Thirty percent of supervisors reported that they preferred some aspects of the form but did not like other aspects. The 30% of supervisors who disliked an aspect of the form used in Method 2 reported that they would have preferred if the form had an

area to write in additional comments for elements of the session that were not specifically targeted.

Results of Research Question 2

Do similarities in a clinical supervisor's background and experience result in greater interrater agreement of assessment and feedback of an observed therapy session? A discussion of the data analysis and findings for Research Question 2 follows.

Data analysis. With respect to the second research question, a statistical analysis could not be conducted due to the small numbers in each subcategory. Using the total number of participants (*N*=5), percentages were calculated to compare subcategories of supervisor's background and experience. Because these percentages were based on a small *N*, the results were interpreted with caution.

The demographic information sheet was used to collect information regarding each participant's background and experience. More specifically, information was collected regarding the number of years as a speech-language pathologist, the number of years as a supervisor, their present employment status, their employment history, and their education. This information was used to answer Research Question 2. Data were analyzed in three ways. First, present employment status was determined, and participants were grouped in three categories: full-time faculty members, clinical supervisors who are also adjunct professors, and clinical supervisors who do not teach any additional course work. Second, their comments were analyzed by determining whether a comment was positive or negative and assigning it to one of the 14 categories. Third, percentages were then calculated for the number of positive and negative comments generated in each category by each of the three subgroups. This procedure was repeated for two additional subgroups, including the number of years as a supervisor and the population (pediatric or geriatric) through whom the participant had gained the majority of their experience.

Findings. As seen in Table 5, three full-time faculty members exhibited 100% interrater agreement in eight categories (activity selection, material selection, modeling, utilizing teaching moments, adapting the session as needed, providing reinforcements, providing feedback, and cuing) in their provision of negative comments when using Method 2. Also using Method 2, they exhibited 100% interrater agreement in one category (task maintenance) in their provision of positive comments. This differed from six adjunct or supervisors and six supervisors who do not teach course work in that neither of these groups exhibited 100% interrater agreement in their provision of negative comments for any of the 14 categories. The participants did, however, exhibit 100% interrater agreement in their provision of positive comments in five of the 14 categories (material selection, modeling, body language, providing feedback, and cuing) when using Method 2. In addition, the full-time faculty members exhibited 100% interrater agreement in their provision of negative comments in three categories (material selection, providing clear instructions, and modeling) when using Method 1. Neither of the other groups reached 100% interrater agreement of either positive or negative comments in any category when using Method 1.

Table 5. Percentage of Positive Comments and Negative Comments for Both Methods 1 and 2 Made by Three Subgroups Based on Employment Status

	Full-time faculty				**Adjunct/ supervisor**				**Supervisor only**			
	Positive comments		**Negative comments**		**Positive comments**		**Negative comments**		**Positive comments**		**Negative comments**	
Category/ Item on form used in Method 2	**M 1**	**M 2**	**M 1**	**M 2**	**M 1**	**M 2**	**M 1**	**M 2**	**M 1**	**M 2**	**M 1**	**M 2**
1. Activity selection was appropriate	0	33	66	100	33	83	33	16	33	100	0	0
2. Material selection was appropriate	33	66	100	100	66	100	33	33	66	100	16	16
3. Clear instructions were provided	33	33	100	66	16	83	0	50	33	83	16	16
4. Modeling was provided	0	0	100	100	50	100	0	0	50	100	16	33
5. Teaching moments were utilized	0	0	66	100	16	83	16	16	33	83	33	0
6. Body language/ enthusiasm	66	66	33	66	16	100	0	33	50	100	0	16
7. Session was adapted to the client's needs	0	66	0	100	0	66	0	0	0	83	0	0
8. Reinforcements were appropriate	0	0	0	100	33	83	0	16	0	50	16	50
9. Feedback was appropriate	0	33	66	100	33	100	0	0	83	83	0	16
10. Data collection was unobtrusive	0	33	0	33	0	50	0	0	0	16	33	33
11. Behavior management was appropriate	0	33	0	0	0	33	0	0	0	50	0	0
12. Session remained on task	0	100	0	0	0	83	0	0	33	83	0	0
13. Cuing was appropriate	0	33	0	100	16	100	0	0	33	100	0	16
14. Assigned homework was appropriate	0	0	0	0	0	16	0	0	0	16	0	16

Note. M = method.

Overall, Method 1 resulted in a total of 128 positive comments and 93 negative comments, and Method 2 resulted in 402 positive comments and 172 negative comments. These numbers illustrated that Method 2 resulted in a greater number of comments in the target areas than Method 1. Method 1 contained comments in 21 different categories. Method

2, although generating more comments overall, contained comments in just the 14 predetermined categories.

In addition, more instances of 100% interrater agreement were noted when Method 2 was used. These results supported the hypothesis that using a structured data collection form with preset areas of clinical performance results in greater interrater agreement among a group of speech-language pathology clinical supervisors.

Discussion

The aim of this study was to determine whether the use of a structured tool that focused on predetermined target areas of a therapy session increased interrater agreement of assessment and feedback among a group of clinical supervisors. This study also explored the interrater agreement of supervisors based on their employment status and clinical and supervisory experience.

A repeated measures ANOVA demonstrated that there was no significant difference between the three videos utilized, no interaction between the videos and the *subjects, and a significant difference between Method 1 (the open-ended data collection form) and Method 2 (the structured data collection form). A series of *t* tests were used to further analyze the results. Three *t* tests were performed: one for each video that was reviewed. Each *t* test revealed greater interrater agreement among the group using Method 2 than using Method 1. In other words, when the group of supervisors viewed the same speech therapy session, they were found to provide more similar positive and negative comments about the clinician's performance when they collected data using the structured form (Method 2) than they did when they collected their data using blank sheets of paper (Method 1). The same result was found for all three videos that were viewed. The *t* tests that were performed confirmed that the use of a structured data collection form with preset target areas of clinician performance resulted in greater interrater agreement among a group of clinical supervisors than an unstructured, open-ended method of data collection. These results have both theoretical and clinical implications.

Theoretical Implications

When using Method 2, the group of supervisors provided positive and negative comments regarding the clinician's performance that were more in agreement with each other than when using Method 1. These findings are in agreement with the findings of Glaser and Donnelly (1989) and Lamb, Orbach, Hershkowitz, Esplin, and Horowitz (2007), who found that use of a structured observational tool when collecting data results in improved quality and less subjective impressions. Kerl, Garcia, McCullough, and Maxwell (2002) stated that a structured tool used for performance evaluations could be helpful as it could reduce subjective judgments. This coincides with Pershey and Reese (2002), who stated that supervisors should select a small amount of behaviors to target for change. Selecting a small amount of behaviors is more manageable when using a structured data collection form with a limited number of clinical behaviors to target than when using an open-ended form that provides unlimited options for supervisor comments. When specific clinical behaviors are predetermined, student clinicians can prepare their sessions accordingly; however, when an open-ended form is used, student clinicians do not know what areas to focus on until after the session is completed.

Further exploration of the comments made by supervisors noted that both methods of data collection contained positive and negative elements in the same category. One example would be, "The materials selected were appropriate, however there was not enough variety of materials." Method 2, however, resulted in a greater number of comments containing both positive and negative elements than did Method 1. Previous research (Ashby & O'Brien, 2007; Boehler et al., 2006; Dowling & Wittkopp, 1982) explored the provision for both positive and negative feedback to student clinicians and found that participants demonstrated improvement of skills when both positive and negative feedback was provided. In addition, Busari, Weggelaar, Knottnerus, Greidanus, and Scherphier (2005) and Vásquez (2004) found that

participants wanted to receive constructive feedback that contained both positive and negative comments. Vásquez found that teaching assistants felt that their program administrator's comments were too polite or too careful in their wording and did not reflect their true opinions when only positive feedback was provided. Busari et al. also reported that attending physicians preferred feedback that was constructive and adapted to the individual's needs. These studies highlighted the importance of supervisory comments containing both positive and negative feedback that is constructive in nature. It is not enough for supervisors to simply say, "You did a good job."

Gillam, Roussos, and Anderson (1990); Lindo (2001); Pershey and Reese (2002); and Shapiro and Anderson (1989) found that selecting specific areas to target for feedback results in greater performance. Kilminster, Cottrell, Grant, and Jolly (2007) stated that learning outcomes should be known by both the supervisor and the supervisee before a task is completed. Selecting predetermined target behaviors for change can be accomplished by using a tool, such as the evaluative statements method.

Clinical Implications

Results of this study revealed practical options for speech-language pathology supervisors to utilize when observing therapy sessions conducted by graduate students in a clinical training program. On the basis of the results of this study, interrater agreement among supervisors proved to increase when using a structured tool. Supervisors suggested that the form should include a blank section for writing in additional comments that are not normally addressed. Adapting this form to create a more concise, one-page form with additional space to write in comments was preferred by the group.

Another clinical implication that is crucial to the development and growth of the graduate clinician is how the form could be adapted from one semester to another. For example, in the beginning weeks of clinical practicum, the form could focus the attention of supervisors and supervisees to only five or six elements, such as specific elements of the session that are often difficult for beginning clinicians, for example, activity and material selection. As clinical practicum experiences expand and those areas are mastered, additional areas could become the focus of the sessions, such as keeping data, task maintenance, and modeling.

This form could be shared with the clinicians so that all parties are aware of the target behaviors, similar to a suggestion presented by Pershey and Reese (2002). The modification of the form would be consistent with Anderson's continuum model of supervision (Anderson, 1988), which stated that there are different strategies that are appropriate during different points in the supervisory process and that supervision exists on a continuum that changes as the supervisee progresses. Anderson's model encourages supervisors and supervisees to develop a relationship that allows the supervisee to grow and become more independent. A dependent relationship between supervisor and supervisee is discouraged. Anderson's model of supervision outlines three stages, including evaluation feedback, transitional, and self-supervision with self-supervision the goal. The form could be adapted to each stage of the continuum, allowing student clinicians to become more independent by perhaps involving them in the selection of clinical behaviors to target for change or to encourage self-assessment. Student clinicians could use the form to assess taped sessions of themselves conducting therapy. Their responses on the form can be compared with those of their supervisor regarding the same session. This type of scenario would follow Anderson's model through the transitional stage and into the self-supervision stage.

One additional benefit to using a structured data collection form is that, following each session, written documentation on the clinician's performance can be provided by the supervisor. Providing immediate feedback has been shown to be more beneficial than providing delayed feedback (Ho & Whitehill, 2009). The use of a form, such as the one utilized in this study, can help to ensure that each student receives immediate feedback following a session even if there is limited time for a conference. This coincides with White's (2008) perspective,

which revealed that written feedback can prove helpful if time prevents a face-to-face meeting. Finally, use of a structured form can serve to provide the university with a written document to support subjective judgments that may not be substantive enough when reviewing possible failure of a student. Students who are required to repeat a clinical practicum due to poor performance can benefit from having a written documentation of their strengths and weaknesses. Weekly written documentation can be kept in the student's file and reviewed, if needed. A technical report developed by the Ad Hoc Committee on Supervision for the ASHA (2008b) revealed that supervisors sometimes encounter students who are considered "marginal" (Supervising Challenging Supervisees section), meaning that they are unable to work independently and follow through on supervisory suggestions. When working with marginalized students, it was recommended in the technical report that supervisors give specific feedback and concrete assistance to the students (ASHA, 2008b). Kerl and colleagues (2002) discussed a similar situation. Kerl and colleagues reported that use of a systematic structured tool for performance evaluations can be helpful as it reduces subjective judgments when the faculty find it necessary to dismiss a student from a program.

The technical report developed by the Ad Hoc Committee on Supervision for the ASHA (2008b) indicated that supervisory training can be obtained in a number of ways, including "course work, continuing education programs, self-study, peer-mentoring, and resources from ASHA" (Training in Supervision section para. 2). Because no formal training procedures are in place across settings, facilities may incorporate a structured method of data collection for supervisors. Training could then be developed specific to that method. The results of this study revealed that the group of supervisors, when thinking divergently when providing free-form feedback, exhibited less interrater agreement than when thinking in a convergent way when using a structured form for feedback.

Dowling (1993) found that supervisors who took part in a supervisory training program met 80 of 87 targeted goals during the supervisory conferences that they held. Supervisors from one facility could be trained to use a structured form for data collection in an effort to increase their ability to meet targeted goals. Hasnain, Onishi, and Elstein (2004) reported that interrater agreement was higher if raters were given specific scoring rules and did not have to rely on their own subjective ideas. Topps, Evans, Thistlethwaite, Tie, and Ellaway (2009) and Shayne et al. (2006) also found that use of predetermined attributes for scoring resulted in greater interrater agreement. The present study supports the findings in the previous studies, which illustrated the benefits of using a structured data collection form when supervising student clinicians in speech-language pathology.

This study also addressed similarities in supervisors' responses based on their employment status and experience. One pattern that emerged was related to the employment status of the participants. It was found that the full-time faculty members produced more negative comments than either the supervisors, adjuncts, or supervisors who did not teach additional course work. In fact, with regard to positive comments, the full-time faculty members reached 100% interrater agreement only once, when reporting that the session remained on task. This proved interesting, as it demonstrated that 12 of 15 supervisors thought that the clinician's choice of materials, use of modeling, use of cuing, and enthusiasm and body language were appropriate for the session, whereas all three full-time faculty members felt that none of these areas was performed appropriately by the clinician. This finding supported those of Geller (2002), who reported that supervisors often turn to their own experiences as a model for how to supervise. Full-time faculty members who have been together at the same facility for a number of years may share similarities in opinions of student performance.

Another interesting finding occurred with regard to the number of years of experience as a supervisor. The group of supervisors with the least experience was the most in agreement with each other. This finding may be the result of supervisors with more experience in the field

having more experiences and events to draw from, whereas supervisors with less experience in the field can draw from fewer experiences and events.

Geller (2002) reported that supervisors who were not given formal training had to rely on their own experiences for how to supervise. Supervisors having more years of experience have more past experiences on which they can rely. Because most facilities, particularly university clinic settings, have supervisors with a large variety of experiences, having a standard form for data collection, and training in the use of that form may prove beneficial for both supervisors and students.

The third area that was analyzed in this study was the interrater agreement among supervisors based on experiences with clinical populations (i.e., pediatrics or geriatrics). Supervisors with either geriatric or pediatric experience achieved greater agreement using Method 2 than they did using Method 1.

Limitations

One limitation to this research was the number of participants. Nelson (2009) stated, "If researchers use sample sizes that are too small, their findings are less likely to provide a true picture of the population as a whole" (p. 135). Huck (2008); Morgan, Reichert, and Harrison (2002); and Schiavetti and Metz (1997) additionally cautioned against using a sample size that is too small; however, Schiavetti and Metz also stated that "between-subject designs usually require larger samples than within-subjects designs" (p. 214).

The research was designed to measure interrater agreement of supervisors at one university; therefore, the number of participants was limited to the number of supervisors at that university. A larger pool of participants could have resulted in the opportunity to further analyze participant responses by state or geographic region. In addition, the results may have potential use with other clinical populations, but generalization to other related disciplines is limited. Extending the recruitment efforts to additional facilities and a larger geographic region could have enlarged the participant pool as well as allowed for a more adequate representation of the characteristics of the population of interest.

Another limiting factor was the length of time necessary to complete both methods. Some potential participants may have declined to participate because of this factor alone. The total time needed to participate was factored into the workday, but scheduling conflicts and other commitments may have prevented some supervisors from participating. Shorter videos would have required less of a time commitment and may have allowed more supervisors to participate. Additionally, supervisors could have viewed the taped sessions at home, thereby eliminating the problem of schedule conflicts and unexpected events the day of the study that prevented some supervisors from participating.

Future Research

Future research should include similar studies at additional facilities. Clinical supervision of graduate students takes place not only at the university setting but also at outside facilities where externships take place. Understanding the feedback provided by supervisors at outside facilities may also prove beneficial to university training programs. This information would also prove helpful during the facility selection process when placing students at these facilities for training. Having a structured form for supervisors at outside facilities to use could improve consistency between facilities and help to provide each student with an equal learning experience. In addition, supervisors at outside facilities may be more willing to accept students for a clinical externship if their data collection and feedback method is streamlined. This could decrease the amount of preparation time they would need in order to provide the student with adequate feedback.

Future research could also explore the student's perception of the feedback that was received. Speech-language pathology graduate students receive feedback from a number of supervisors over multiple semesters. Understanding students' perceptions of that feedback and

understanding what feedback is viewed as most beneficial and least beneficial would be helpful for university training programs.

Technology used in supervision could also be explored in future research. Advances in technology integrated with supervision have led to concepts such as remote supervision and electronic conferencing. This application could be designed for data collection, with supervisors completing data collection forms on an iPad or other device that could be immediately shared with the student being observed.

Conclusions

This study aimed to answer two questions: first, whether the use of a structured tool that focused on predetermined target areas of a therapy session increased interrater agreement of assessment and feedback among a group of clinical supervisors, and, second, whether similarities in a clinical supervisor's background and experience resulted in greater interrater agreement of assessment and feedback of an observed therapy session. The results of this study indicated that there was a statistically significant difference between the interrater agreement of supervisors when one used the structured data collection form as opposed to the recording evaluative statements method of data collection; interrater agreement was greater when the supervisors used the structured tool. Additionally, it was determined that full-time faculty members generally provided more negative comments than supervisors who were employed as adjunct faculty members and supervisors who do not teach any course work. It was also determined that supervisors with less supervisory experience achieved greater interrater agreement than supervisors with more years of supervisory experience. These findings indicated that additional research and supervisory training could be beneficial to clinical supervisors.

Clinical supervision creates a "triangulated" (Chabon, Hale, & Wark, 2008) relationship among the supervisor, student, and client. Supervisors play an essential role in both advancement of the student's skills as a clinician and the client's skills in therapy. Often, supervisors must balance the ethical considerations presented in this dynamic. On a daily basis, supervisors must weigh the best interest of the student with the best interest of the client and determine a course of action. Often, these decisions are difficult given the subjective nature of the field. In many situations, a formal, objective measurement would aid in the supervisory process. It is impossible to eliminate subjective opinions from supervisory feedback; however, providing supervisors with training and formal performance evaluation tools will impact both student clinicians and clients alike and provide them both with invaluable skills for the future.

Acknowledgments

This article is based on a doctoral dissertation completed at Nova Southeastern University under the direction and guidance of Dr. Angela Sherman.

Comments/questions about this article? Visit our ASHA Community and join the discussion!

References

American Speech-Language Hearing Association. (2008a). *Clinical supervision in speech-language pathology* [Position statement]. Retrieved from http://www.asha .org/docs/html/PS2008-00295.html.

American Speech-Language-Hearing Association. (2008b). *Clinical supervision in speech-language pathology* [Technical report]. Retrieved from www.asha.org/ policy.

Anderson, J. L. (1988). *The supervisory process in speech-language pathology and audiology*. Austin, TX: Pro-Ed.

Ashby, F., & O'Brien, J. (2007). The effects of positive versus negative feedback on information-integration category learning. *Perception & Psychophysics, 69,* 865–878.

Bischoff, R., Barton, M., Thober, J., & Hawley, R. (2002). Events and experiences impacting the development of clinical self-confidence: A study of the first year of client contact. *Journal of Marital and Family Therapy, 28,* 371–382.

Boehler, M., Rogers, D., Schwind, C., Mayforth, R., Quin, J., Williams, R., & Dunnington, G. (2006). An investigation of medical student reactions to feedback: A randomised controlled trial. *Medical Education, 40,* 746–749. doi:10.1111/j.1365-2929.2006
.02503.x

Busari, J., Weggelaar, N., Knottnerus, A., Greidanus, P-M., & Scherphier, A. (2005). How medical residents perceive the quality of supervision provided by attending doctors in the clinical setting. *Medical Education, 39,* 696–703. doi:10.1111/j
.1365-2929.2005.02190.x

Chabon, S. S., Hale, S. T., & Wark, D. J. (2008, February 12). Triangulated ethics: The patient-student-supervisor relationship. *ASHA Leader*. Retrieved from http://www
.asha.org/Publications/leader/2008/080212/080212f/.

Dowling, S. (1993). Supervisory training, objective setting and grade-contingent performance. *Language, Speech, and Hearing Services in Schools, 24,* 92–99.

Dowling, S., & Wittkopp, J. (1982). Students' perceived supervisory needs. *Journal of Communication Disorders, 15,* 319–328. doi:10.1016/0021-9924(82)90014-4

Geller, E. (2002). A reflective model of supervision in speech-language pathology: Process and practice. *Clinical Supervisor, 20,* 191–200. doi:10.1300/
J001v20n02_14

Gillam, R., Roussos, C., & Anderson, J. (1990). Facilitating changes in supervisees' clinical behaviors: An experimental investigation of supervisory effectiveness. *Journal of Speech and Hearing Disorders, 55,* 729–739.

Glaser, A., & Donnelly, C. (1989). Data-based supervision methods for speech-language pathology. *Language, Speech, and Hearing in Schools, 20,* 296–304.

Hasnain, M., Onishi, H., & Elstein, A. (2004). Interrater agreement in judging errors in diagnostic reasoning. *Medical Education, 38,* 609–616.

Ho, D. W. L., & Whitehill, T. (2009). Clinical supervision of speech-language pathology students: Comparison of two models of feedback. *International Journal of Speech-Language Pathology, 11,* 244–255.

Hoben, K., Varley, R., & Cox, R. (2007). Clinical reasoning skills of speech and language therapy students. *International Journal of Language and Communication Disorders, 42,* 123–135.

Huck, S. W. (2008). *Reading statistics and research.* Boston, MA: Pearson Education.

Indiana University. (1996). *Indiana University evaluation of speech-language pathology practicum (IUESLPP).* Bloomington: Author, Department of Speech and Hearing Sciences.

Irwin, D. L., Pannbacker, M., & Lass, N. J. (2008). *Clinical research methods in speech-language pathology and audiology.* San Diego, CA: Plural.

Kerl, S. B., Garcia, J. L., McCullough, S., & Maxwell, M. E. (2002). Systematic evaluation of professional performance: Legally supported procedure and process. *Counselor Education and Supervision, 41,* 321–334.

Kilminster, S., Cottrell, D., Grant, J., & Jolly, B. (2007). AMEE Guide No. 27: Effective educational and clinical supervision. *Medical Teacher, 29,* 2–19.

Lamb, M., Orbach, Y., Hershkowitz, I., Esplin, P., & Horowitz, D. (2007). A structured forensic interview protocol improves the quality and informativeness of investigative interviews with children: A review of research using the NICHD investigative interview protocol. *Child Abuse & Neglect, 31,* 1201–1231.

Lindo, D. (2001, September 1). Produce information--not martininfoo! *SuperVision, 62,* 7–10.

London, M. (2003). *Job feedback: Giving, seeking and using feedback for performance improvement* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

McCrea, E., & Brasseur, J. (2003). *The supervisory process in speech-language pathology and audiology*. Boston, MA: Pearson Education.

Morgan, S. E., Reichert, T., & Harrison, T. R. (2002). *From numbers to words*. Boston, MA: Pearson Education.

Nelson, L. K. (2009). *Research in communication sciences and disorders*. San Diego, CA: Plural.

Pershey, M., & Reese, S. (2002). Consumer satisfaction with speech-language pathology services in university clinics: Implications for student supervision. *Clinical Supervisor, 21,* 185–205.

Schiavetti, N., & Metz, D. E. (1997). *Evaluating research in communicative disorders.* Needham Heights, MA: Macmillan.

Shapiro, D., & Anderson, J. (1989). One measure of supervisory effectiveness in speech-language pathology and audiology. *Journal of Speech and Hearing Disorders, 54,* 549–557.

Shayne, P., Gallahue, F., Rinnert, S., Anderson, C. L., Hern, G., & Katz, E. (2006). Reliability of a core competency checklist assessment in the emergency department: The standardized direct observation assessment tool. *Academic Emergency Medicine, 13,* 727–732.

Shriberg, L., Filley, F., Hayes, D., Kwiatkowski, J., Shatz, J., Simmons, K., & Smith, M. E. (1974). *The Wisconsin procedure for appraisal of clinical competence (W-PACC).* Madison: University of Wisconsin–Madison, Department of Communicative Disorders.

Smith, H. (1984). Moment-to-moment counseling process feedback using a dual-channel audiotape recording. *Counselor Education and Supervision, 23,* 346–349.

Topps, D., Evans, R. J., Thistlethwaite, J. E., Tie, R. N., & Ellaway, R. H. (2009). The one minute mentor: A pilot study assessing medical students' and residents' professional behaviors through recordings of clinical perceptors' immediate feedback. *Education for Health: Change in Learning and Practice, 22,* 1–11. Retrieved from EBSCO Open Access Journals database.

Vásquez, C. (2004). "Very carefully managed": Advice and suggestions in post-observation meetings. *Linguistics and Education: An International Research Journal, 15,* 33–58.

White, S. (2008). Using action research to gauge the quality of feedback given to student teachers while on professional teaching practice. *Qualitative Research Journal, 8,* 124–133.