

Relationship Development Between Racially Matched and Non-matched Counselor Supervisors and Practicum Supervisees: Preliminary Findings.

James H. Payne, South University, Savannah
Shon D. Smith, Argosy University, Sarasota
Barry Tuchfeld, Traditional Aikido of Sarasota
Joffrey S. Suprina, Argosy University, Sarasota

Abstract

Results of racially (RM) and non-racially matched (NRM) supervisory pairs where supervisors attempted to predict supervisees' ratings of the supervisory working alliance (SWA), using the Leeds Alliance in Supervision Scale (LASS; Wainwright, 2010) were explored. The Total LASS Dyad Difference Scores (i.e., difference between the supervisee's rating and supervisor's prediction of the rating) showed a statistically significant relationship in the point biserial correlations at Week 1; there were no statistically significant relationships at Weeks 3 or 5 for the RM and NRM dyads. With the small sample size these results must be considered exploratory.

Bordin (1983) described the working alliance as one involving building a strong alliance and then repairing ruptures, which he called the rupture-repair process. A rupture was where the collaborative bond resulted in conflict or a breakdown in the therapeutic working alliance. Bordin (1979, 1983) reported that there were three factors involved in the working alliance: goals, tasks, and bonds. Chen and Bernstein (2000) found that the bond/relationship between the supervisor and the supervisee must be established first as it cannot withstand a challenge prior to the establishment of the bond.

Successful establishment of a bond may not be sufficient to address challenges of non-racially matched supervisory pairs. Riley (2004) found that perceptions of Caucasian supervisors' do not necessarily match that of their supervisees of Color. In fact, she found an inverse relationship between supervisees' ratings on a multicultural instrument and supervisors rating of the supervisory working alliance (SWA). She concluded, "mixed race supervision dyads are vulnerable to perceptual differences" (p. 85). Supervisees of Color were defined as African-American, Hispanic American, or Asian American.

To address perceptual differences, Milne (2009) described supervision as "Culturally (an) effective practice (that) necessitates a rare degree of self-awareness and sensitivity, including the ability to respond appropriately to the individual characteristics of one's supervisees" (p. 105). Dressel, Consoli, Kim, and Atkinson (2007) found four skills that may lead to "unsuccessful multicultural supervision" (p. 62): lack of sensitivity, inflexibility, rigidity, and dogmatism. Gatmon, Jackson, Koshkarian, Martos-Perry, Molina, Patel, and Rodolfa (2001) found, "providing an atmosphere of safety, depth dialogue, and frequent opportunities to discuss

cultural variables in the supervisory relationship contributed to building alliances and increasing satisfaction” (p. 108). Gatmon et al. reported that a cultural match between supervisee and supervisor was not necessary for the alliance to be satisfactory. In fact, it was “the presence and quality of the discussion of difference and similarities ... (that made) supervisors ... better equipped to initiate (discussions about multicultural issues)” (pp. 110-111). Such open discussion may support the quality improvement in the supervisory relationship that Ladany, Mori, and Mehr (2013) reported was an important ingredient in effective supervision. However, Hess, Knox, Schultz, Hill, Sloan, Brandit et al. (2008) reported that interns in effective supervisory relationships did not initiate certain discussions unless first initiated by their supervisors. Thus, supervisors must provide a safe place where they initiate all relevant discussions, including those about differences in the supervisory relationship. In addition, initiation by the supervisor of discussions and differences may lead to more open communications.

Initiation of the discussion will benefit from emotional sensitivity. Torres-Rivera, Wilbur, Maddux, Smaby, Phan, and Roberts-Wilber (2002) found that emotional sensitivity was a necessary ingredient in the training and development of counselors’ multicultural skills and competencies. They suggested measurement and assessment of the SWA. As in counseling, assessment and measurement of the SWA has an impact on supervision outcome. Supervision has a mandated component for all trainees as students, as well as, prior to full licensure. While supervision was mandated, the quality of this process has never received as much emphasis.

Assessment may help explain any variance as Ladany, Brittan-Powell, and Pannu (1997) concluded, “Racial identity interaction may account for a significant amount of unexplained variance in the supervision process” (p. 301). They suggested that it was not about racial matching alone, or even level of racial identity development; this impacted cultural competencies and level of racial identity develop and its interaction on the dyad pairs. The supervisor must be as open to the process as the trainee/student, for supervision to work effectively.

Although Inman (2006) found a statistically significant positive correlation between supervisor multicultural competence and the SWA which led him to suggest “the working alliance may be a significant “common factor” in multicultural supervision and needs to be integrated into the conceptualization of multicultural competence in supervision” (p. 83), Riley (2004) found that “when supervisors perceived themselves as having good interpersonal relationships with their multicultural clients, their trainees of color perceived their working alliance as less positive” (p. 84). She reported the importance of discussion of multicultural issues in supervision. This supports openness to multicultural discussions as a necessary component of multicultural competence.

The necessity of multicultural discussion is supported by Nilsson and Dobbs (2006), who found two factors involved in working effectively with international students: Multicultural Discussion (46% of the outcome variance) and Supervisees’ Cultural Knowledge (12% of the variance). When supervisees bring up cultural issues in supervision it created a strain in the SWA, which was reported as “perceived prejudice” (p. 226). This perceived prejudice ... was significantly correlated with both role ambiguity and role conflict, indicating that more experiences of prejudice were associated with more uncertainty regarding

supervisor's expectations and evaluations and how to manage the sometimes contradictory roles of being a student, supervisee, colleague, and counselor simultaneously. ... supervisory relationships are not isolated from the social contexts in which we live and that racial and ethnic minority students experiences of perceived prejudice, among other factors, are associated with their experiences in supervision. (p. 226)

In supervision, the supervisor carries both the power and ability to resolve and initiate discussions of similarities and differences, including discussion of multicultural similarities and differences, which supports the findings of Ladany et al. (1997).

Such discussions are also beneficial with gender variance. Walker, Ladany, and Pate-Carolan (2007) found that when working with female supervisees, supervisors who engaged in discussion of gender issues agreed more on the tasks of supervision. All supervisors regardless of their gender who did not engage in these discussions had lower supervisee ratings on the SWA. Assertive supervisory power resulted in weaker agreement on the goal and task scales, along with a weaker emotional bond. As in research in other multicultural areas (Allen, 2007; Dressel et al., 2007; Inman, 2006; Nilsson & Dobbs, 2006; Nilsson & Duan, 2007), and supervisory power (Nelson & Friedlander, 2001; Murphy & Wright, 2005), it was the supervisor's openness that invited discussion, thus increasing the likelihood of effective processing of conflict.

To access the impact of multicultural discussions on the SWA, effective assessment instruments are needed. Watkins (1997) stated, "one of the most pressing needs for psychotherapy supervision in the next century remains the development and establishment of reliable, valid criterion measures to guide supervision" (p. 94). Early instruments were instruments designed for use in counseling were simply converted to use in supervision without the requisite validity, reliability, and norms for use in supervision. In 2007, Torres-Rivera et al. supported Watkins' earlier call, which encouraged Wainwright's (2010) quest to develop such an instrument.

Wainwright (2010) reviewed existing instruments designed to measure the SWA and concluded that they were too time consuming for use on a regular basis. Using a combination of qualitative and quantitative methods he developed the three items of the Leeds Alliance in Supervision Scale (LASS) labeled: Approach, Relationship, and Meeting my needs. The Approach subscale measured the degree of structure and focus during a supervisory session. The Relationship subscale measured how the supervisee perceived the quality of the supervisory relationship. The Meeting my needs subscale measured how the supervisor helped the supervisee set, clarify, and meet goals. This resulted in a brief (i.e., under one minute) supervision sessional rating scale. The LASS provided a score for each item and a total score, which was the sum of the three scores whose range was 0 to 300. Although brief, Wainwright pointed out that the LASS was similar to other analogue measures of the working alliance in counseling/psychotherapy (e.g., the Session Rating Scale, Duncan, Miller, Sparks, Claud, Reynolds, et al., 2003).

Methods

Participants

Practicum students and their campus supervisors (i.e., these were a combination of faculty and doctoral students supervising master clinical mental health counseling students) were requested to participate. These were from 10 Council for Accreditation of Counseling and Related Educational Programs (CACREP) accredited programs in the Southeastern United States, with two programs agreeing to participate (20%). The two programs provided a total of 32 dyad pairs: Institution A with 25 pairs supervised by 11 doctoral student interns (the doctoral students were enrolled in a CACREP doctoral program); and Institution B with 7 pairs supervised by 2 doctoral level licensed professionals (a CACREP mental health counseling program).

Supervisors' demographic summary was as follows:

1. Gender: Females 11/13 (84.6%) and males 2/13 (15.4%);
2. Age: Average age 43, standard deviation (*SD*) of 11.98 with a range from 27 to 64;
3. Campus supervisors: 2/13 were doctoral level licensed professionals in CACREP clinical mental health counseling program the and remaining 11/13 supervisors had master's degree, serving as supervisors as part of their doctoral training (84.6%); and 2/13 from one campus, both licensed with doctorates (15.4%);
4. Racial identity: 4/13 as Black/African American (30.8%), 8/13 as Caucasian (61.5%); and 1/13 as Hispanic (7.7%);
5. Professional experience as a licensed counselor: $M = 86.17$ months, $SD = 58$ with a range of 14 to 184 months;
6. Supervisory experience: $M = 32.87$ months, $SD = 52.91$ with a range of 0 to 196 months;
7. Professional licensure: 7/13 licensed in their state as a counselor (53.8%), 1/13 (dually licensed in psychology and marriage and family counseling) (7.7%), and 5/13 registered counseling interns/equivalents (38.5%);
8. State approved supervisors: 6/13 were (46.2%), and 7/13 were not (53.8%); and
9. Two of the 13 supervisees had previous experience with the supervisors as course instructors (15.4%), with the remaining (11/13) 84.6% without previous experience with their supervisees.

Supervisee demographic data was as follows:

1. Gender: 30/32 females (93.7%) with 2/32 males (6.2%);
2. Age: $M = 34.56$, $SD = 11.2$, range 24 through 69;
3. Campus: Institution A: 25/32 (78.1%) and Institution B 7/32 (21.9%);
4. Racial identity: Bi-racial/multiracial 3/32 (9.4%), Black/African American 10/32 (31.3%), Caucasian 17/32 (53.1%), and Hispanic 2/32 (6.2%);
5. Prior experience in working in a mental health setting: 15/32 identified prior experience (46.9%), with 17/32 as no prior experience (53.1%);
6. Total academic hours (i.e., semester hours towards their degree): 26 participants (81.2%) reported their total number of academic hours (semester) at the beginning of the semester, $M = 46.62$ (semester hours) with a SD of 14.22, with a range from 9 to 80 hours – both programs were 60 semester hour clinical mental health counseling programs;
7. All supervisory dyads had a supervision contract – 32/32 (100%); and

8. Match on racial identity: 15/32 matched (46.88%) with 17/32 did not match (53.12%) (See Table 1).

Table 1
Racial Match of Supervisee/Supervisor

		Supervisor			
		Blk/AA	Cau	Hispanic	Bi/Multi
Supervisee	Blk/AA ¹	3 ¹	8		
	Cau	6	10	1	
	Latin		1	2	
	Bi/Multi	1			

¹One supervisee identified as bi-racial and said there was a match with the supervisor, although the supervisor identified as Black/African American – and was therefore identified as a racial match (RM).

Procedures

The LASS was administered by campus supervisors to their supervisees in individual/triadic supervision at Week 1 of the research project, then again at Weeks 3 and 5. The supervisor asked the supervisee to rate the session just completed (Weeks 1, 3, and 5); independently the supervisor predicted the supervisee’s ratings on the LASS. At Weeks 2 and 4 the supervisor began the session by asking the supervisee to discuss their ratings from the prior week, followed by how s/he predicted the supervisee rated the session. They discussed their similarities and differences. The total discussion was limited to a maximum of five minutes. There was no formal research question as the findings were part of a larger study (Payne, 2012).

The data split almost evenly, racial match (RM) of 15 supervisees and supervisors along with nonracial match (NRM) of 17. Table 2 provides the descriptive statistics for the two groups at Weeks 1, 3, and 5. The LASS Total Dyad Difference Scores (LTDDS) take the supervisee’s rating of the SWA for each week less the supervisor’s prediction for that session. These are mean difference scores. The subscale scores are so highly intercorrelated (Wainwright, 2010) necessitating the use of the LASS Total Dyad Difference Scores.

Table 2

Statistics for Racial Match (RM) and Non-Racial Match (NRM) LTDDS

	Week 1	Week 3	Week 5
Mean Racial Match (RM) N = 15	18.2667	2.300	2.333
Mean No Racial Match (NRM) N = 17	-7.971	1.441	4.882
Standard Deviation RM	27.217	23.284	11.841
Standard Deviation NRM	40.280	25.247	12.703
Point biserial correlation ($n = 32$) $r_{biserial}$.362	.018	-.106
$p =$.042	.921	.563

Results

Inspection of the descriptive statistics for Week 1 (Table 2) showed there was a difference between how supervisee and supervisors rated this session. A positive mean score indicated that the supervisees rated the session higher than the supervisors, while a negative mean score indicated the supervisors rated the session higher. A point biserial correlation was done (Howell, 2010), where one of the variables was categorical (i.e., Racial Matched, Non-racial Match) and the other was interval or ratio (i.e., LASS Total Dyad Difference Scores). The point biserial correlation for Week 1 was ($r = .362, n = 32, p = .042$) meaning that beginning with Week 1 there was a small positive correlation between the two groups, $R^2 = .1310$, accounting for 13.1% of the variance. For Week 3 the biserial correlation was not significantly correlated ($r = .018, n = 32, p = .921$). In Table 2 this was evident by the decrease in difference in the mean scores between the two groups (RM = 2.300 and NRM = 1.441); the reader will see that the standard deviations decreased from Weeks 1 to 3 as well. A similar trend continued into Week 5 ($r = -.106, n = 32, p = .563$) with a continuing decrease in standard deviations for both groups.

The beginning relationship may have been due to the differences between the supervisees' rating the sessions higher than the supervisors for the racially matched group and the supervisors rating the SWA higher than the supervisees for the non-rationally matched group, resulting in the negative scores. There was no qualitative data describing the content of any of the sessions for Week 1, 2, 3, 4, or 5. The observation was that Caucasian supervisors rated the sessions higher than supervisees in the NRM at Week 1.

Total LASS Scores of supervisees improved from Weeks 1 to 2, Weeks 1 to 5, and Weeks 2 to 5 (Figure 3 and Table 3). SPSS 20 was used for the analysis of all the data. A pre-analysis of the LASS Total Scores for supervisees was done to see if they met multivariate assumptions. There were a total of 3 outliers initially; an interactive process continued until only one outlier remained leaving a total of 28 supervisees. There was no significance in Mauchleys test of sphericity ($W = .904, \chi^2 = 2.634, df = 2, p = .268$). A repeated measures analysis (Warner, 2008) was done with LASS Totals over 3 administrations ($F = 4.574, df = 2/54, p = .015$; and $\eta^2 = .145$) (See Figure 3 LASS Total Means by Weeks for Supervisees). Post hoc analyses found there was no significant difference between Weeks 1 and 3 ($t = -.869, df = 27, p = .392$). There was a significant difference between Weeks 3 and 5 ($t = -2.663, df = 32, p = .012$), also, a significant difference between Weeks 1 and 5 ($t = -3.390, df = 27, p = .003$) (See Table 4 Post-hoc Analysis). (See Table 3 for a summary of the descriptive statistics).

Table 3
LASS Total Statistics for Supervisees

	Week 1	Week 3	Week 5
Mean	278.607	281.607	287.574
Standard Deviation	13.557	16.807	11.487
Variance	183.792	282.475	131.951
Skewness	-.723	-1.049	-.263
Kurtosis	.294	.544	-1.278

Figure 3
LASS Total Means by Weeks for Supervisees

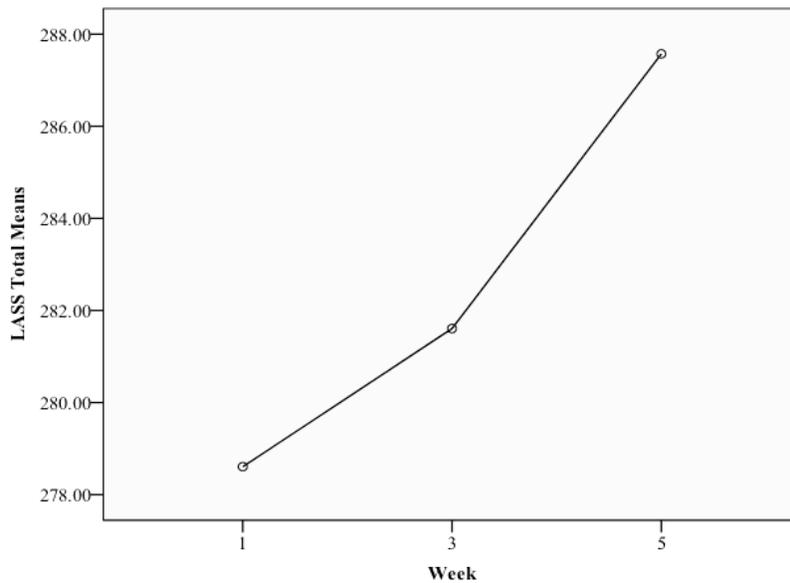


Table 4
 Post-hoc Analysis

	<i>t</i>	<i>df</i>	<i>p</i>
Weeks 1 to 3	-.869	27	.392
Weeks 1 to 5	-2.663	32	.012
Weeks 3 to 5	-3.390	27	.003

Discussion

Riley (2004) and Ladany et al. (1997) collected their data during a singular session. This study used a time series design early in the semester experience of practicum students. The authors of this study looked at the course of the development of supervisors' ability to predict the ratings of supervisees' ratings of the SWA using the LASS. It also looked at the development of SWA from the supervisees' perspective. Overall, supervisors improved their predictive ability whether there was a racial match (RM) or non-racial match (NRM). The predictive ability of supervisors improved over the five weeks of the study; there was a discussion of similarities and differences from the beginning for all pairs. Bachelor and Horvath (1999) found that counselors were unable to accurately predict clients' ratings of the therapeutic working alliance. While the working alliance has determined to have differences between counseling and supervision; there was a similarity in predictive ability of counselors and supervisors early in the processes. This similarity may be a principle in the development of most alliances. Lampropoulos (2001) described predictive ability as learned rather than intuitive.

The split between racial and non-rationally matched pairs was not part of the original research design. At Week 1 supervisees rated the total working alliance higher than the supervisors by an average total of 18.27 for the RM group. Thus, on average the supervisor

predicted a lower Total LASS Score than did their supervisee. At Week 1 the NRM group supervisees rated the sessions lower than the supervisors predicted. Analysis of the results supports that there were different strategies employed by supervisors in the RM as compared with that of supervisors in the NRM group. Over the period of the five weeks supervisors improved their predictive abilities in both of these groups. There was no clear rationale for the initial discrepancies. With multicultural training there has been an emphasis upon both fairness and accuracy. The RM supervisees appeared to have rated the session based on fairness, while supervisors rated it on accuracy. This resulted in an average of 18.26 points difference (positive numbers represent the supervisees rating the session higher than the supervisor's predicted). Conversely, the NRM supervisees appeared to have rated it based on accuracy, while the supervisors predicted based on fairness. The supervisees rated the sessions lower than supervisors predicted (i.e., an average of -7.971 differences). This correlation at Week 1 was much lower by Week 3 and remained so at Week 5. Without a reference point, the strategy was based on either/or (i.e., either fairness or accuracy) depending on the initial match between the pairs. With experience, thus a reference point, the participants shifted to a both/and strategy involving both fairness and accuracy. While this study found a relationship between RM and NRM at Week 1, this relationship was no longer evident by the Week 3 and remained afterwards consistent through Week 5. Supervisors' predictive ability adapted as they developed a stronger SWA with their supervisees. This improvement occurred quickly with general instructions to discuss similarities and differences. The results added general support to the literature review rather than specific support. Supervisors improved their predictive ability with supervisees, whether in the RM or NRM by Weeks 3 and 5.

The beginning differences resulted from supervisors' lack of experience with the supervisees. The differences in predictive ability between the RM and NRM groups shifted from Week 1 to Week 3 as a result of a general discussion of similarities and differences. Supervisors of both groups continued to improve their predictive ability in Weeks 3 and 5. The instructions regarding discussions were open-ended rather than specific to encourage dyad communication strategies. Each week supervisors began the discussion regarding similarities and differences based on the ratings from the previous supervision session. Results were based on a general discussion of similarities and differences, rather than on specific directions.

Equally significant was supervisees' continued progressive ratings of the SWA over Weeks 3 and 5 (see Figure 3 and Table 3). This was evidence for the developmental nature of SWA. Supervisors helped manage this through improving their ability to predict supervisees' ratings of the SWA. Scores were based on the Total LASS Scores rather than the individual subscales.

The results provide support for regular use of assessment in supervision. Traditionally, assessment in supervision was limited to research studies. These findings support the recommendations made Watkins (1997) and Torres-Rivera et al. (2007) regarding regular assessment in supervision. The use of assessment for this research design made it possible to capture the results of discussion about similarities and differences during the supervision process; had this not been done the alternative would have been unclear. The assessment process also encouraged discussions that may have otherwise been neglected.

The sample size limited the external validity of this study, thus it remains an exploratory study. At the beginning (Week 1) there was a correlation within RM and not in the NRM pairs. This relationship disappeared by the third week when the LASS was again administered. Thus the dyads became more comfortable with each other and racial matching appeared to be of less importance, which may have resulted from the open discussions of the discrepancy in LASS scores. Further research would determine the exact change mechanism.

The general findings from this exploratory study found that while racial matching did impact the first session and there was a relationship with racially matched pairs, this beginning relationship disappeared sometime between Weeks 2 and 3 (i.e., Week 3 was the second time the LASS was administered). What was found was that the Total LASS Scores improved over time as rated by supervisees; in addition, the differences between the actual ratings of supervisees and supervisors' predictions of supervisees ratings was lowered (i.e., LASS Total Dyad Difference Scores got smaller indicating a greater match of actual and predicted). This indicated the establishment of bonds, goals, and tasks of the SWA, thus the relationship was maintained through Weeks 4 and 5. Even during this time supervisees continued to rate the SWA at levels of greater satisfaction. Higher SWA scores by supervisees related to higher satisfaction in supervision. The findings must be cautiously considered important; further research will either confirm or disconfirm these findings. The researchers predict that like Bachelor and Horvath (1997), predictive ability of counselors and supervisors will be a skill that will improve with practice. Practice must include on-going or periodic assessment and measurement.

In this study a disproportionately higher number of female supervisees and supervisors were evident. Future research would benefit from researching how this impacts supervision in same-sexed pairs, as well as opposite sexed pairs; it would be interesting to discover if there was a similar relationship at Week 1 as there was in RN and NRM pairs. The impact of measurement upon the process must also be a consideration in future research.

References

- Allen, J. (2007). A multicultural assessments supervision model to guide research and practice. *Professional Psychology, 38*, 248-258.
- Bachelor, A. & Horvath, A. (1999). The therapeutic relationship. In M. A. Hubble, B. L. Duncan, & S. D. Miller (Eds) *The Heart & Soul of Change: What Works in Therapy* (pp. 133-178). Washington, DC: American Psychological Association.
- Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16*, 252-260.
- Bordin, E. S. (1983). A working alliance based model of supervision. *The Counseling Psychologist, 11*, 35-42.
- Chen, E. C. & Bernstein, B. L. (2000). Relations of complementarity and supervisory issues to supervisory working alliance: A comparative analysis of two cases. *Journal of Counseling Psychology, 47*, 485-497.
- Dressel, J. L., Consoli, A. J., Kim, S. K., & Atkinson, D. R. (2007). Successful and unsuccessful multicultural supervisory behaviors: A Delphi poll. *Journal of Multicultural Counseling and Development, 35*, 51-64.

- Duncan, B. L., Miller, S. D., Sparks, J. A., Claud, D. A., Reynolds, L. R. et al. (2003). The Session Rating Scale: Preliminary psychometric properties of a "working" alliance measure. *The Journal of Brief Therapy, 3*, 3-12.
- Gatmon, D., Jackson, D., Koshkarian, L., Martos-Perry N., Molina, A., Patel, N., & Rodolfa, E. (2001). Exploring ethnic, gender, and sexual orientation variables in supervision: Do they really matter? *Journal of Multicultural Counseling and Development, 29*, 102-113.
- Hess, S. A., Knox, S., Schultz, J. M., Hill, C. E., Sloan, L., Brandit, S., et al. (2008). Predoctoral interns' nondisclosure in supervision. *Psychotherapy Research, 18*, 400-411.
- Howell, D. C. (2010). *Statistical Methods for Psychology* (7th ed.). Belmont, CA: Cengage Wadsworth.
- Inman, A. G. (2006). Supervisor multicultural competence and its relation to supervisory process and outcome. *Journal of Marital and Family Therapy, 32*, 73-85.
- Ladany, N., Mori, Y., & Mehr, K. E. (2013). Effective and ineffective supervision. *The Counseling Psychologist, 41*, 28-47.
- Ladany, N., Brittan-Powell, C. S., & Pannu, R. K. (1997). The influence of supervisory racial identity interaction and racial matching on supervisory working alliance and supervisee multicultural competence. *Counselor Education and Supervision, 36*, 284-304.
- Lampropoulos, G. K. (2001). Common process of change in psychotherapy and seven other social interactions. *British Journal of Guidance & Counseling, 29*, 21-33.
- Milne, D. (2009). *Evidenced-based Clinical Supervision: Principles and Practices*. West Sussex, UK: Wiley.
- Murphy, M. J. & Wright, D. W. (2005). Supervisees' perspectives of power use in supervision. *Journal of Marital & Family Therapy, 31*, 283-295.
- Nelson, M. L. & Friedlander, M. L. (2001). A close look at conflictual supervisory relationships: The trainee's perspective. *Journal of Counseling Psychology, 48*, 384-395.
- Nilsson, J. E. & Dodds, A. K. (2006). A pilot phase in the development of the International Student Supervision Scale. *Journal of Multicultural Counseling & Development, 34*, 50-62.
- Nilsson, J. E. & Duan, C. (2007). Experiences of prejudice, role difficulties, and counseling self-efficacy among U.S. racial and ethnic minority supervisees working with white supervisors. *Journal of Multicultural Counseling & Development, 35*, 219-229.
- Payne, J. H. (2012). *Supervisor predictions of supervisee ratings of the supervisory working alliance utilizing the Leeds Alliance in Supervision Scale (LASS)* (Doctoral Dissertation).
- Riley, S. M. (2004). *An investigation of the relationship between counseling supervisors' multicultural counseling competencies and the supervisory working alliance* (Doctoral Dissertation). UMI Number: 3123395.
- Tores-Rivera, E., Wilbur, M. P., Maddux, C. D., Smaby, M. H., Phan, L. T., & Roberts-Wilbur, J. Factor structure and construct validity of the counseling skills personal development rating form. *Counselor Education & Supervision, 41*, 266-278.
- Wainwright, N. A. (2010). The development of the Leeds Alliance in Supervision Scale (LASS): A brief sessional measure of the supervisory alliance. (Published doctoral dissertation). Retrieved from <http://etheses.whiterose.ac.uk/1118/>. (ID Code: 1118).
- Walker, J. A., Ladany, N., & Pate-Carolan, L. M. (2007). Gender-related events in psychotherapy supervision: Female trainee perspectives. *Counseling & Psychotherapy Research, 71*, 12-18.

Warner, R. M. (2008). *Applied Statistics*. Thousand Oaks, CA: Sage.

Watkins, C. E. (1998). Psychotherapy supervision in the 21st century. *Journal of Psychotherapy Practice & Research*, 7(2), 94-101.