Impact of Two Clinical Peer Supervision Models on Practicing School Counselors

By: Lori B. Crutchfield, L. DiAnne Borders


Abstract:

Pre- and posttest measures were used to assess school counselors' level of job satisfaction, counseling self-efficacy and counseling effectiveness. None of the analyses of covariance examining treatment effects were significant.

Keywords: Counseling | Schools | Job Satisfaction | Psychology

Article:

Two forms of clinical peer supervision were provided for a sample of 29 practicing school counselors. Pre- and posttest measures were used to assess the counselors' level of job satisfaction, counseling self-efficacy, and counseling effectiveness (including empathic responding, adaptability and flexibility in counselor response, and client behavior change). The sample was divided into 3 groups (2 treatment and 1 control). Each supervision treatment lasted 9 weeks. None of the analyses of covariance examining treatment effects were significant. However, these individually nonsignificant results showed movement in the hypothesized direction in each instance, indicating small but pervasive effects of treatment. Participants' qualitative session evaluations also supported the helpfulness of clinical peer supervision for school counselors. Implications for future research and the practice of school counselor supervision are discussed.

Within the public schools, the person best trained to help children combat their problems while continuing their healthy development is the school counselor. Well versed in counseling skills and techniques, school counselors work with children individually and in small groups, and they provide large-group guidance lessons when needed (American School Counselor Association, 1993). Despite their skills in these areas, however, it is becoming increasingly difficult for school counselors to provide adequate counseling services for the diverse counseling issues and student groups represented in their schools.

There are several reasons for this increasing difficulty. First, whereas most mental health counselors work in settings with other counselors present, many elementary and middle school counselors are isolated in their settings (Peace, 1995). In fact, the elementary school counselor often serves two or more schools on an itinerant basis. Second, unlike mental health counselors, school counselors usually receive little or no consistent counseling supervision (American...
Consequently, school counselors can become unsure of their counseling abilities, even possibly becoming less skilled than they were upon receiving their counseling degrees (Peace, 1995; Spooner & Stone, 1977). Lack of sufficient supervisory support also increases stress and intensifies immense workloads. Thus, school counselors may feel overworked, alone, burnt out, and unhappy with their roles. As a result, counselors may become less effective, and therefore less able, to provide meaningful help to the children in their schools. Third, in some cases, school counselors may have completed training before many skills needed to address today's problems (e.g., family counseling, suicide assessments) were offered in training programs. Although ongoing professional development is encouraged for school counselors by the American School Counselor Association (1993), there is rarely a systematic means available for providing such continued training.

One way to provide needed support for school counselors is through clinical supervision. A structured though varying set of activities that encourages counselor self-awareness and growth, clinical supervision can focus on skill enhancement, professional identity development, case conceptualization, or other aspects of the school counselor's role in providing direct service to young clients (e.g., Bernard & Goodyear, 1992). There is empirical support for the theoretical need (developmental models; Blocher, 1983; Hogan, 1964; Loganbill, Hardy, & Delworth, 1982; Sansbury, 1982; Stoltenberg, 1981; see Worthington, 1987, for a review of the literature) and the desire of counselors themselves to receive clinical supervision (Borders & Usher, 1992; Roberts & Borders, 1994).

The impact of clinical supervision on counselors' performance was illustrated by Wiley and Ray's (1986) study. These researchers concluded that counselors benefited more from supervised experience and showed little growth from unsupervised experience. Henderson and Lampe (1992) emphasized the personalized nature of counseling supervision as an enhancer of school counselor professional development. The school counselor who engages in supervised clinical experiences is likely to make positive professional changes that will lead to greater effectiveness and accountability (Borders, 1991). By providing feedback and questions from an objective third party, clinical supervision allows the counselor to view the client and the counseling situation from multiple perspectives, thus stimulating the counselor's integrative thinking about the case (Biggs, 1988).

Even in the 1970s, the lack of supervision for school counselors was noted. Boyd and Walter (1975), for example, compared the school counselor to a cactus, saying that, by necessity, both must grow and thrive with the minimal amount of "nutrients" (p. 103). Since the 1970s, many other authors have expressed similar concerns about the almost nonexistent supervision of school counselors (American Association of Counseling and Development Task Force, 1989; Barret & Schmidt, 1986; Borders, 1991; Borders & Schmidt, 1992; Schmidt, 1990; Schmidt & Barret, 1983; Wilson & Remley, 1987). In fact, what is called supervision for school counselors is most often administrative oversight being supplied by school principals (American Association of

In a national survey, Borders and Usher (1992) found that school counselors were receiving the least supervision of counselors in a variety of settings. Similarly, Roberts and Borders (1994) found that, although school counselors spent 44 of their time in counseling and consultation, they received little or no supervision in these areas. These findings demonstrate the immense discrepancy between existing supervision practices and school counselors' needs for supervision.

Although the lack of supervision is often attributed to school systems' limited funding for trained supervisors (e.g., Schmidt & Barret, 1983), it also is due in part to school administrators' increasing concerns about counselor time spent in direct service to students. Clinical supervision may be seen as a less-than-useful reason for taking school counselors away from this direct service. Admittedly, providing individual clinical supervision for school counselors may be an inappropriate goal, because of both time and budget constraints within school systems.

Various suggestions for time- and cost-effective approaches to supervision have been made in the literature (e.g., Benshoff & Paisley, 1996; Borders, 1991b; Fraleigh & Buchheimer, 1969; Henderson & Lampe, 1992; Hillerbrand, 1989; Holloway & Johnston, 1985; Lewis, Greenburg, & Hatch, 1988; Peace, 1995; Remley, Benshoff, & Mowbray, 1987; Roth, 1986; Spice & Spice, 1976; VanZandt & Perry, 1992; Wagner & Smith, 1979). In general, these approaches address restrictive administrative needs within public schools through the use of mentors, peer dyads, and peer groups.

Peer-group consultation and supervision—less threatening approaches to self-examination and professional growth than supervision within a hierarchy—are becoming increasingly popular means of efficiently providing clinical supervision to practitioners (Benshoff & Paisley, 1996; Borders, 1991b; Remley et al., 1987; Runkel & Hackney, 1982). Whether in dyads or small groups, peers can provide one another support and encouragement, as well as the challenge to think about their clients in new ways. Just as group counseling is more time efficient than individual counseling when working with students, group supervision is an effective means of providing clinical supervision for a number of counselors concurrently (Borders, 1991b; Fraleigh & Buchheimer, 1969; Greenburg, Lewis, & Johnson, 1985; Hillerbrand, 1989; Holloway & Johnston, 1985; Lewis et al., 1988). Thus, peer groups can provide support and encouragement, as well as enhance skills and promote personal and professional development (Lewis et al., 1988; Yalom, 1985). In addition, dyadic peer interaction provides collegial support and an egalitarian setting in which school counselors can interact in productive exchanges about professional issues.

Recently, Benshoff and Paisley (1996) offered one model based on peer interaction designed specifically for practicing school counselors. An adaptation of the Remley et al. (1987) model, the Structured Peer Consultation Model for School Counselors (SPCM-SC) provides clearly
structured activities to promote egalitarian support and professional growth and development. This dyadic model was tested with a small sample of school counselors in North Carolina. Overall, the participating school counselors reported that they had positive consultation and supervision experiences during the study. Counselors agreed that they would participate in peer consultation and supervision again and would recommend it to their colleagues. They also reported that participation in the model had helped them develop better consultation skills, as well as improve their counseling skills and techniques (Benshoff & Paisley, 1996). Finally, 100 of the participants believed that the peer consultation and supervision had provided them with valuable support and new ideas, and that it had been a helpful experience. Most believed that the experience would have been less valuable without the structure provided.

Similarly, Borders (1991b) delineated a model of structured peer-group supervision designed to promote skill development as well as conceptual growth. She agreed with others (Roth, 1986; Runkel & Hackney, 1982) that, without such structure, peers often stray from the supervisory task or give each other advice rather than productive feedback. Borders’ (1991b) model, reported to be appropriate for both counselors-in-training and experienced practitioners, includes strategies for promoting cognitive skills development. According to numerous authors (e.g., Biggs, 1988; Blocher, 1983; Borders, Fong, & Neimeyer, 1986; Holloway, 1988), higher functioning counselors (those at higher cognitive developmental levels) are more likely to think independently, objectively, and flexibly and to express empathy with a greater variety of clients. This flexibility and empathy could translate into greater counseling effectiveness by providing the counselor with a larger repertoire of counseling approaches and techniques. Within the Borders (1991b) model, divergent thinking is promoted through the assignment of diverse roles, multiple theoretical approaches, and creative metaphors.

In summary, there are several strong statements supporting clinical supervision of school counselors, suggesting that such supervision would help them grow professionally and personally. It also appears that peer-group approaches with some structure might be the supervision mode of choice. To date, however, few peer or peer-group models have been implemented, and even fewer evaluated for their impact. Thus, the purpose of this study was to investigate the assumption that peer-group clinical supervision can have a positive impact on the effectiveness of school counselors. Two forms of clinical peer supervision (one dyadic and one group) were provided for practicing school counselors, and the impact of these interventions on job satisfaction, counselor self-efficacy, and counseling effectiveness was measured. To determine whether clinical peer supervision makes a difference for practicing school counselors, we used a pretest-posttest design to assess the counselors’ level of each variable both before and after the supervision was provided. More specifically, we investigated the following research questions:

1. Does clinical peer supervision have a positive effect on practicing school counselors' job satisfaction perceptions?
2. Does clinical peer supervision have a positive effect on practicing school counselors' perceptions of counseling self-efficacy?

3. Does clinical peer supervision have a positive effect on practicing school counselors' counseling effectiveness?

4. Which of the two models of clinical peer supervision is most helpful to practicing school counselors?

METHOD

Participants

Twenty-nine practicing school counselors from a rural area in the Southeast volunteered to participate. Eight counselors were assigned to the first (dyadic) treatment group, 10 to the second (peer-group) treatment group, and 11 to the unstructured (control) group. Participants were assigned to groups upon consideration of several factors relevant to the feasibility of group membership. Because of the wide geographic region represented by the counselors, they first were grouped by area (e.g., counties). Also, some of the counselors had previous experience with one of the two treatment models; these counselors were assigned to the control group. No counselor in either treatment group had postdegree experience with either of the supervision models. Finally, because of the nature of the two treatment groups, counties with larger numbers of participants were more likely to be assigned to the peer-group supervision model. The final number of participants in the first treatment group, originally 10, fell to 8 when one of the pairs of participants withdrew from the study.

Each participant received a small stipend, with those in the treatment groups ($125 each) being paid more than those in the control group ($25 each). As those in the control group only met twice (compared with 10 meetings for those in the treatment groups), they received the lesser amount. Most participants were elementary and middle school counselors, with two high school counselors participating. The majority (83) had a master's degree as the highest degree earned. Twenty-four were female, and all were White. Ages ranged from 25 years to 56 years, with the majority (69) of participants falling in the 33- to 46 year range.

Peer Supervision Interventions

The first treatment group participated in the SPCM-SC (Benshoff & Paisley, 1996). This group was trained in the model of dyadic peer consultation-supervision by a trained, experienced supervisor familiar with the model. Meeting once a week, for 9 weeks, participants in this group provided supportive yet challenging peer consultation-supervision to their partners following an adapted structured protocol. Peer consultation-supervision sessions included setting individual goals, review and discussion of partner's counseling session audiotapes, and case presentations and discussions.
Training of participants consisted of a general introduction to the model (Benshoff & Paisley, 1996), followed by a videotaped demonstration of its use within a consultation-supervision session. A training manual, created in consultation with the model's authors, was distributed to participants, and all questions were addressed. During the weeks of this intervention, the only checks on implementation were weekly Post Session Helpfulness Questionnaires (PSHQ; Hill, 1989), which participants completed and immediately mailed to the researcher. This showed that the dyads were meeting on a weekly basis, but no systematic checks were made to discern how closely participants were implementing the SPCM-SC.

The second treatment group participated in Systematic Peer Group Supervision (SPGS; Borders, 1991b). This model uses systematic assignments of particular roles (e.g., counselor, student, and student's teacher) within the group of supervisees during audiotape reviews. After the presenting counselor's taped segment has been heard, each group member responds to the counselor's prestated questions in the first person, from the role they were assigned. Listening and responding to tapes from these roles encourages conceptualization and skill-building participation by each of the members. The supervisor facilitates the process by asking specific questions of the different roles, asking the counselor to respond, and summing up the statements and suggestions of the group.

Training of participants consisted of a brief general introduction of the model (Borders, 1991b), followed by a videotaped demonstration of its use within a supervision session. A training manual, created in consultation with the model's author, was distributed to participants, and all questions were addressed. During the 9 weeks of group supervision (meeting once a week), the large group of 10 participants met in two smaller groups of 5, each led by a trained supervisor experienced in the model. The supervisors for each group reported following the SPGS model closely within each session.

The third group served as the unstructured "pseudotreatment" (control) group and completed the pre- and posttest battery of questionnaires only. Members of this group were asked to focus individually on their plans for professional development during the time of the study, meeting only twice (at the beginning and the end of the study). The data gathered from this group provided a comparison point for the two treatment groups.

Dependent Variables

Job satisfaction. To measure job satisfaction, we used an adapted version of the Job Satisfaction Blank (JSB; Hoppock, 1977). The JSB consists of four items (e.g., "Choose one of the following statements which best tells how well you like your job") in a Likert-scale format (e.g., 1 = "I hate it" and 7 = "I love it"). A global score of job satisfaction is derived by summing the weighted responses to the four items, with higher scores suggesting higher global job satisfaction. Scores can range from 4 to 28, with high satisfaction designated as scores of 23 and up; average satisfaction scores designated as 16-22, and low satisfaction scores as 15 and below (Wiggins &
Moody, 1983). Hoppock (1977) reported split-half reliability as $r = .87$, and Brayfield and Roethe (1951) found a correlation of .67 between the JSB and the composite score of 257 questions regarding conditions of work and job satisfaction. The JSB is not only a brief, reliable, and valid measure (McNichols, Stahl, & Manley, 1978); it also has been used successfully with samples of school counselors (Gade & Houdek, 1993; Wiggins & Weslander, 1986).

Self-efficacy. Participants indicated individualized judgment of their capacity to perform satisfactorily in given counseling situations on the Counseling Self-Estimate Inventory (COSE; Larson et al., 1992). The self-report, 37-item questionnaire uses a 6-point Likert-type rating scale, ranging from strongly disagree (1) to strongly agree (6). The highest possible final score (item score sum) is 222; higher scores indicate higher self-estimates of counseling efficacy (Larson et al., 1992). This instrument is the first general measure of counseling self-efficacy to be developed, and the only one with established validity and reliability information. A factor analysis (N = 213) yielded five factors around which the 37 retained items were clustered: Microskills (alpha = .88), Process (alpha = .87), Difficult Client Behaviors (alpha = .80), Cultural Competence (alpha = .78), and Awareness of Values (alpha = .62). Computed internal consistency for the total inventory was alpha = .93. A short form of the COSE (COSE-SF), with 30 items, correlates highly (.99) with the COSE total score and has 3-week test-retest coefficients of .87 for COSE-SF total, .83 for Awareness of Values, .80 for Difficult Client Behaviors, .74 for Process, .71 for Cultural Competence, and .68 for Microskills (Larson et al., 1992). For the purposes of this study, COSE total scores served as the pretest-posttest comparison points.

Counseling effectiveness. Three instruments were used to measure different aspects of counseling effectiveness. The first aspect, that of empathic responding, was measured by the Index of Responding Empathy Scale (IRE; Gazda et al., 1984a). This 10-item scale consists of helpee statements to which participants are asked to respond empathically, writing out the empathic response directly below the helpee statement. For the purposes of this study, minor changes in wording were made to the measure, so that the participant responded as the counselor (vs. teacher) in all cases, and the helpee was identified either as a student/client or a parent or teacher consultee.

Responses are rated according to Gazda and associates' (Gazda et al., 1984b) 4-point Empathy scale. A response that would be rated at Level 1.0 on the scale is considered irrelevant to the helpee's statement; it is also hurtful to the helpee, because it does not attend even to the surface feelings involved. A Level 2.0 response communicates the content of the helpee's statement accurately but is still considered subtractive because it only partially attends to the surface feelings. A response at Level 3.0 reflects the helpee's surface feelings adequately, and, if content is included, it is accurate. A Level 4.0 response is considered additive because the helpee's underlying feelings are identified, and content may be used to add a deeper meaning (Gazda et al., 1984b). If raters supply different scores, a consensus on the rating is required.
Test-retest reliability coefficients ranging from .90 to .92 have been reported (Black & Phillips, 1982; Gazda et al., 1984b) for this measure. Gazda et al. also reported split-half reliability as .77. Internal consistency was computed as alpha = .76 (Cummings & Murray, 1990). In addition, numerous experts agree that the instrument holds a great deal of face validity (Black & Phillips, 1982; Cummings, 1989; Cummings & Murray, 1990; Davis et al., 1985; Hector et al., 1981). In the present study, trained raters used a manual developed by Cummings (1989). This manual contains specific examples of responses at each possible level. Training consisted of at least 10 hours of discussion and practice using written materials and responses produced by individuals not participating in this study. Once an acceptable level of interrater reliability (.70) was obtained, training was complete. Final interrater reliability ranged from .57 (on pretest, agreement between Rater 1 and Rater 2) to .79 (on posttest, overall agreement). The mean interrater reliability (for pre- and posttest) was .67. It is generally agreed that .70 is the minimum acceptable level of reliability; this is discussed further in the Discussion section.

The second aspect of counseling effectiveness, flexibility of response, was measured by the Counselor Behavior Analysis Scale (CBA-Long; Howard, Nance, & Myers, 1987), a 24-item self-report measure of counselor adaptability. The first 12 items from the CBA-Long make up the CBAA, and the second 12 items make up the CBA-B. According to Adaptive Counseling and Therapy theory (Howard, Nance, & Myers, 1986), the basis for the CBA, counselors respond from four different developmental levels, which are on a continuum from low maturity to high maturity. High maturity (adaptable) counselors are better able to match their clients' developmental level within therapy, thus maximizing the therapeutic process. That is, if a client is functioning at a concrete cognitive developmental level, the adaptable counselor would be most likely to recognize this and use an appropriate form of intervention (such as behavior modification). The instrument is reported to be "content-free" (G. S. Howard, personal communication, October 12, 1994), making it appropriate for counselors in all settings.

When psychometric data for the CBA were explored (Gabbard, Howard, & Dunfee, 1986), internal consistency was demonstrated (N = 44) by computing part-whole correlations for the CBA-A and the CBA-B scores with CBA-Long score. For the CBA-A, r = .78, whereas r = .77 for the CBA-B. Gabbard et al. reported a median test-retest reliability coefficient (kappa) of .60 across 3 months for the CBA-Long total score. In addition, counselor adaptability ratings correlated highly with overall counseling effectiveness (.98). For the purposes of this study, only the 12 items on the CBA-B were administered. In an informal pre-pilot study, a group of school counseling experts found the CBA-Long to be too time-consuming and tiring. These experts agreed that Form B alone would be less intimidating to school counselors. According to Howard, both Form A and Form B can stand alone as a valid measure of adaptability (G. S. Howard, personal communication, October 12, 1994).

Client change, the third aspect of counseling effectiveness, was measured by the Teacher Report Form (TRF; Achenbach, 1991), a standardized measure of teacher judgment regarding students' adaptive functioning and problems in school. The instrument consists of a demographic sheet,
ratings of academic performance, comments on four ratings of adaptive functioning, and 118 possible problems classified into eight syndrome scales. With these syndrome scales, the problem scales may be separated into two groupings: Internalizing (e.g., somatic complaints, anxieties) and Externalizing (e.g., delinquent behavior, aggression) problems.

When completing the TRF, teachers use a Likert-type scale, ranging from 0 (not true [as far as you know]) to 2 (very true or often true). In scoring the TRF, we computed total raw scores for each scale by adding the Is and 2s entered for that scale. A graphic display of raw scores allows the use of both normed percentiles and T scores for each scale. In the present study, raw scores were compared with normed percentiles to place them below, within, or above the normal range.

In the instrument manual, Achenbach (1991) reported test-retest reliability (N = 44) over a mean of 15 days of .90 for academic and adaptive scores and .92 for problem scores. Content validity was established by comparing scores of mental health or special education referred students with those of nonreferred students (N = 2,600; see Achenbach, 1991). Referred students scored significantly higher (p <.01) on nearly all of the problem score items, suggesting that the TRF items are related to mental health needs. Achenbach (1991) established construct validity by a comparison of TRF scores and scores on the Conners Revised Teacher Rating Scale (Goyette, Conners, & Ulrich, 1978). The correlation between the TRF and Conners Total Problems scales was .83.

Supervision session helpfulness. For exploratory purposes only, the Post-Session Helpfulness Questionnaire, an adaptation of Hill's (1989) Client Post-Session Questionnaire, was used as a measure of the helpfulness of the supervision sessions. This measure is made up of three items. On the first item, participants are asked to rate the supervision session on the basis of a 5-point Likert-scale, where 1 = not helpful and 5 = extremely helpful. The remaining items are open-ended questions aimed at gathering general feedback about possible helpful and harmful elements of each session.

Procedure

We assigned participants to one of three groups after taking into consideration several factors relevant to the feasibility of group membership (e.g., geographic location and previous supervision experience). At the pretest meetings, groups of participants were given the pretest packet of instruments and were asked to complete the measures in the order in which they appeared (as listed above). Participants were encouraged to answer as honestly as possible and were reminded that all information would be kept confidential, with only identification numbers used on the instruments. Completion of the packets took approximately 1 to 1.5 hours. At this point, participants were asked to set up their schedules for the supervision sessions.

After the pretest and scheduling were completed, the use of the TRF (Achenbach, 1991) was explained and participants each received 10 copies of the instrument. For the purposes of this study, participating school counselors had teachers of the first five clients they saw during the
study complete the demographic information and the problems checklist of the TRF (Achenbach, 1991) as soon as possible after the first counseling session. Counselors were asked to choose students whom they expected to see at least three times over the course of the study. During the final week of the study, this instrument was completed again by the teachers of each of these clients who was seen at least three times by the participating school counselors. Counselors also asked teachers to complete TRFs (at pre- and posttest) for 1 control student per client. This student was the 12th student on the teacher's roll, in the same class as the client. Counselors were asked to match the students by gender; thus, if the 12th student was not the same gender as the client, the 13th student (or the next student who was the appropriate gender) was used. It was left up to the participating counselors to decide who the control student would be, then ask the teacher to complete the forms; the teacher remained unaware of the designations of client and control students.

RESULTS

A series of one-way analyses of covariance (ANCOVAs) were used to test for significant treatment effects on each of the dependent counselor variables (job satisfaction, perceived counseling self-efficacy, and counseling effectiveness [measured by level of empathic responding and counselor adoptability in responding]) in comparison with the control group. A three-way analysis of variance was used to test for significant treatment effects on the dependent variable of client behavior change.

Descriptive Results

Pretest and posttest scores on each of the instruments were calculated for participants. All scores were plotted for each measure and the distributions appeared normal. Means and standard deviations are reported in Table 1. In general, pretest results revealed average job satisfaction (JSB) scores (M = 20.9). Self-efficacy pretest scores (COSE) were relatively high (M = 168). For the written measure of empathy (IRE), pretest scores were low enough to be considered "subtractive" (M = 2.2). Adaptability pretest scores on the CBA-B were slightly above average (M = 35.6).

At posttest, job satisfaction (JSB) scores again were average (M = 21.2), though slightly higher than at pretest. Self-efficacy posttest scores (COSE) were still relatively high (M = 172.3), again slightly higher than at pretest. On the written measure of empathy (IRE), posttest scores again were subtractive (M = 2.4), though again slightly higher than at pretest. Adaptability scores on the CBA-B were still slightly above average (M = 36.4), also rising slightly from pretest.

Analysis of TRF scores is based on gender and age. (See Table 2 for means and standard deviations.) For this group of students, overall raw scores on the TRF were high-normal, meaning that, in general, students had raw scores within the range of normal behavior, though at the high end of this range. Scores above the cutoff for the normal range imply the presence of clinical behavioral problems. For the boys, the TRF Internalizing subscale pretest mean (8.06),
Externalizing subscale pretest mean (13.18), and total TRF mean at pretest (40.27) were all at the highest end of the normal range, indicating that their behavioral problems were near clinical proportions. Boys' posttest means (Internalizing = 7.14, Externalizing = 13.49, and Total = 40.12) were nearly identical to means at pretest.

For the girls, the TRF Internalizing subscale pretest mean (8.92) and total TRF pretest mean (31.64) were, much like the boys', at the highest end of the normal range. However, on the Externalizing subscale, their pretest mean (10.12) was firmly within the clinical range. Girls' posttest means (Internalizing = 6.46, Externalizing = 8.36, and Total = 25.15) were all slightly lower than at pretest, though the Externalizing mean was still barely within the clinical range. (TRF scores are separated by gender for clients and controls below.)

Scores for all the measures also were calculated by treatment group. Means and standard deviations by treatment group are presented in Tables 3 and 4. For the counselors' dependent measures at pretest, groups were similar, as they were at posttest.

Regarding the client information, results by treatment groups were again similar. The small differences that did occur were between the two groups of students (clients and controls). On the pretest, student clients scored significantly higher on all scales than did control students Internalizing, $F(2,157) = 36.26$, $p = .0001$; Externalizing, $F(2, 145) = 26.93$, $p = .0001$; and Total, $F(2, 145) = 53.27$, $p = .0001$-as might reasonably be expected, given that the students had been referred to the counselor (higher scores indicate more problematic behaviors). No significant differences were noted for posttest results by counselor treatment group, but control students did achieve significantly greater gains on the posttest (resulting in significantly lower problem behavior scores) than did clients, $F(2,145) = 7.03$, $p = .0089$. There were also significant differences by gender, with girls showing significantly lower posttest scores compared with boys, $F(2,145) = 9.29$, $p = .0027$.

Tests for Treatment Effects

To test for treatment effects, we performed a one-way ANCOVA on posttest scores, with pretest scores used as the covariate, on each of the dependent measures (the counselor assessments), using an overall .05 alpha level. For the JSB, there was no significant main effect by treatment group, $F(2, 25) = 0.90$, $p = .4201$, nor was there any significant main effect by treatment group for the COSE, $F(2, 25) = 0.11$, $p = .8953$. Neither of the other two counselor measures showed significant main effects by treatment group: for the IRE, $F(2, 25) = 0.25$, $p = .7809$, and for the CBA-B, $F(2, 25) = 1.34$, $p = .2808$. Thus, neither supervision group had a significant effect on school counselors' job satisfaction, self-efficacy, or counseling effectiveness.

For the measure of client change (TRF), a three-way ANCOVA on posttest scores with pretest scores, client gender, and client groups used as covariates was performed. On the Internalizing scale (signifying problem behaviors such as depression), there was no three-way interaction effect, $F(2, 145) = 0.43$, $p = .6529$, and no two-way interaction effects for treatment group by
student group, \( F(2,145) = 1.38, p = .2543 \), or by student gender, \( F(2, 145) = 0.32, p = .7248 \). There also was no significant main effect for treatment group, \( F(2, 145) = 1.06, p = .3508 \). On the Externalizing scale (signifying problem behaviors such as delinquency), there was no three-way interaction effect, \( F(2, 145) = 1.12, p = .3296 \), and no two-way interaction effects for treatment group by student group, \( F(2,145) = 1.54, p = .2180 \), or by student gender, \( F(2,145) = 0.28, p = .7534 \). There also was no significant main effect for treatment group, \( F(2, 145) = 0.49, p = .6131 \). And finally, on the Total scale, there was no three-way interaction effect, \( F(2,145) = 0.31, p = .7361 \), and no two-way interaction effects for treatment group by student group, \( F(2,145) = 0.41, p = .6655 \), or by student gender, \( F(2,145) = 0.47, p = .6269 \). There also was no significant main effect for treatment group, \( F(2, 145) = 0.85, p = .4274 \). No significant differences in treatment effects were found, thus no student group (client or control), gender, or treatment group had TRF scores that were significantly different from any other group or groups.

In summary, none of the ANCOVAs examining treatment effects were significant. Thus, there was no significant improvement in school counselors' job satisfaction, counseling self-efficacy, or counseling effectiveness. However, these individually nonsignificant results showed movement in the preferred direction each time, indicating small but pervasive effects of treatment.

**Postsession Questionnaires**

The exploratory PSHQs, completed after each session by counselors in the two treatment groups, provided qualitative evaluations of the supervision sessions. On the first item on the questionnaire, based on a 5-point Likert scale, participants rated the level of helpfulness of the session. Means and standard deviations were calculated for this item by treatment group, per session and overall (see Table 5). The means, for each group and overall, were all above 4, indicating that supervision sessions, regardless of modality, were seen as relatively helpful.

There was great consistency in the responses to PSHQ Questions 2 and 3 regarding the most helpful and most hurtful aspects of the supervision session. Approximately 90 indicated that the feedback and the support received were most helpful. Comments about helpful feedback typically referred to suggestions and ideas offered by other counselors, as well as new insights achieved as a result of discussing cases or listening to tapes. Responses concerning support expressed interest in a continuing dialogue with other counseling professionals, with comments such as "It's good to know there are other counselors struggling with similar issues." Regarding the most harmful aspect of the sessions, most participants (70) stated that nothing was harmful. Those who did answer this item (30) provided concrete responses that often did not really address the question. Example responses were, "It was difficult to hear the tape," "We didn't have enough time to cover everything we needed to talk about," "People came in the room and interrupted us," "I had a hard time getting parental permission to tape," and "Scheduling a time to meet was tough."
DISCUSSION

Results of the study indicated that, over a brief period of time (approximately 2.5 months), clinical peer supervision did not have a statistically significant impact on job satisfaction, self-efficacy, or counseling effectiveness of school counselors participating in the study. When a basic sign test was used, however, small gains on each measure were consistently noted. That is, when change did occur from pretest to posttest, the scores moved in a positive (or desired) direction. This would seem to indicate that the two supervision treatments did have a slightly positive impact on each of the counselor-related dependent variables. Nonetheless, scores on job satisfaction remained in the average range, whereas counseling self-efficacy scores remained fairly high. These results seem to support previous researchers' findings and conclusions regarding the overall stability of job satisfaction and self-efficacy as constructs in adults (Bandura, 1977, 1982; Hoppock, 1977; Smith, Kendall, & Hulin, 1969; Wiggins & Weslander, 1986).

Counseling effectiveness scores as measured by adaptability in response style remained somewhat above average. However, counseling effectiveness scores as measured by empathic responding remained low. In fact, as an aggregate, the school counselors tended to problem-solve instead of responding empathically, which was the main reason for their receiving subtractive ratings (2.0-level scores) on the IRE. Instead of clearly reflecting the feelings the client expressed to them, participants often responded something like this: "I'd like to try to help you figure out a way to resolve that dilemma. What are some things you could do to begin improving the situation?"

Finally, clients maintained significantly higher scores (indicating increases in undesirable behavior) than did controls across the brief span of time covered in the study. This could well be an instance of problem behaviors escalating before they improve, although the brevity of the study did not allow an investigation of this possibility. The significant difference found between genders (boys consistently scored significantly higher than girls on the problem behavior scales) supports the norm set of the instrument (Achenbach, 1991).

In line with the consistent positive trends on several of the measures (i.e., JSB, COSE, IRE, and CBA), participants' subjective responses (on the PSHQ) supported the idea that clinical peer supervision is helpful to school counselors. Counselors' written comments indicated the value of discussing clients and approaches, as well as the helpfulness of role-plays in viewing the situation from different perspectives (in the peer group). In debriefing sessions after the posttest administrations, counselors made similar positive comments. Some examples of verbal responses include: "This type of interaction with other counselors improves my ability to assess and evaluate students," "It helped me look at the whole client," "I realized I have been really isolated, and it's been good to meet regularly with another counselor," "The role-plays helped renew my enthusiasm for the profession," and "It provided validation that what I'm doing is important."
In these qualitative data, two major themes appeared. Participants felt they had gained from the interventions in the areas of (a) collegial/professional support and (b) concrete feedback on counseling skills, approaches, and perspective taking. However, empirical results indicated that no significant changes were made. There are three possible explanations for these conflicting results. First, it is possible that the peer supervision interventions do not actually work to increase counseling effectiveness. This was the first known attempt to actually measure changes in counseling effectiveness resulting from the two peer supervision interventions, as both models were previously untested in terms of their theoretical outcomes. Perhaps, with the dyadic model, there was too much inconsistency without a trained supervisor present to provide structure. It was assumed that the counselors using this model knew what to do during their sessions (because they had been trained) and that they followed through with the established protocol (although there were no systematic checks to ensure this). There may not be enough quality control of skill learning within this model for its use to encourage development in this area. With the peer-group model, the greater awareness levels of differing perspectives and skill enhancement experienced within the group sessions may not have been easily generalized to the school setting.

Second, it is possible that the measures used in this study were either inappropriate for school counselors or did not measure variables or aspects of the counselors' effectiveness or roles that did change. For example, no measure of case conceptualization was used, a cognitive skill implied by several of the counselors in the debriefing sessions as an area of growth. In addition, the measures may not have been sensitive enough to measure subtle beginnings of positive developmental change exhibited by the participating counselors.

A third possible explanation is that the interventions were too brief. It may be that more time was needed to affect change, particularly for counseling practitioners who have had little or no counseling supervision since graduating from their counselor training programs (0-21 years for participants in this study). Benshoff and Paisley (1996) implemented one of the treatments over an entire semester. No pre/post comparison measures were reported (school counselors only completed a general measure of satisfaction with and evaluation of the experience); as a result, no data are available to determine if this longer intervention period was more effective. It should be noted, however, that participants in the Benshoff and Paisley study indicated a desire for a much longer period (at least a year) of peer consultation and supervision before they thought lasting professional change would occur. It may well be, then, that with a brief intervention such as the one used in the present study, small but pervasive effects in a positive direction are the best one can hope for.

Several limitations of the design of this study undoubtedly affected the results. Limitations related to the sample are evidence of the numerous difficulties encountered in conducting field-based experimental studies (Borders, 1989). The sample consisted solely of volunteers who came from a restricted, although quite large, geographical area in North Carolina. Not only did this create a limited sample; it also had quite an impact on assignment to treatment groups. Participants typically needed to travel 30 to 90 minutes to arrive in a central location for
meetings. Because of these geographical restrictions, true random assignment to experimental groups was not feasible. In addition, the small sample size contributed to low statistical power. All of these sampling issues limit the generalizability of the results, as well as the types of data analysis possible and the ability to detect any real changes.

Another limitation was found within the instruments. All measures in this study involved self-report, which can be a questionable measure of any variable (Borders, 1989). The measure of empathic responding required written responses to printed client situations, which may not be as valuable as responding verbally to observed client behaviors. On the measure of flexibility in response styles, the counselor was asked not to choose the ideal response, but the one she or he was most likely to use. The TRF relies on teachers' self-reports of their perceptions of student behaviors. If the teacher has developed a set of expectations for a certain student's behavior, a change of student behavior might not be reported even though it had definitely occurred. Teacher reports, however, seemed less problematic than other measures of outcomes we might have used, such as student reports of satisfaction with counseling sessions (particularly those of elementary students; see Cobb & Richards, 1983; Ritchie, 1989).

Keeping in mind that the interrater reliability mean of .67 fell just short of the commonly accepted .70 minimum, ratings of empathic responses (nonhelpful range) are still somewhat alarming. Do school counselors lack ability in this most basic common therapeutic component (Frank, 1982; Grencavage & Norcross, 1991) and one of Rogers' (1957) core conditions? It may be that, without ongoing supervision, these school counselors had regressed in their ability to perform basic counseling skills. This explanation would be in line with Spooner and Stone's (1977) results, indicating that, without consistent supervision, counselors experienced stagnation or regression in the counseling skills taught in their training programs. If such regression has and does-occur, there are dire implications for the counseling effectiveness of consistently unsupervised school counselors.

It is not clear, however, how representative these low empathic scores are of all school counselors. In a small pilot study involving school counselors from another rural area in North Carolina (Crutchfield, 1995), similarly low empathic responses were noted. In addition, these results seem in line with school counselors' requests for help with specific skills, as reported in two surveys of supervision needs and preferences (Borders & Usher, 1992; Roberts & Borders, 1994). Nevertheless, additional studies are needed to verify the actual counseling performance of school counselors on measures of empathy and other counseling skills.

Admittedly, the school counseling environment is one that encourages "quick fixes" for students seeking counseling. As such, the possible press within the school for counselors to problem solve rather than to reflect feelings might explain participants' IRE responses. It also may be that the school counselors were using a brief solution-focused approach to therapy, which would suggest that a different measure of counseling effectiveness is needed when working with this population.
During the final meetings with the school counselors in the treatment groups, the researcher requested general impressions of the sessions. Without fail, each participant enthusiastically reported how helpful the experience had been. Many expressed a desire to continue meeting, perhaps once a month instead of once a week. There was a sense of rejuvenation and collegial support among the participants, and most were sorry to see the supervision sessions end.

Implications

An examination of the PSHQ uncovered an interesting trend. In general, the school counselors participating in peer consultation dyads reported support to be the most helpful thing about the sessions, whereas school counselors in the peer groups reported specific feedback on skills and techniques to be most helpful. This evidence suggests that, in the absence of a trained supervisor, the focus is on collegial support. With a trained supervisor present, it may be more likely that skill development and enhancement can occur, although not in a brief time period such as in this study. Perhaps some combination of the two models would be a more effective approach to fulfill the supervision needs of school counselors. If peers met in dyads weekly, then in a group once a month, there may be more consistency in the structure provided, as well as the more appropriate balance of challenge and support needed to produce growth and development.

If this topic is researched in the future, it would be best to engage in a long-term study (6 months to 1 year). This attempt to find an impact on counseling effectiveness over a brief period was fruitless, but not futile. It appears that more time may be needed for supervision interventions to affect counselors and their work, especially those who have had little clinical supervision in their careers. A true random sample might be possible within large school systems, particularly if the counseling supervisor at the central office level was willing to set up an experimental clinical supervision program for all the school counselors in the system over a year's time. In addition, continued use of multiple measures of counseling effectiveness is supported, given the range of scores obtained on the three instruments. Any replication of this research would be strengthened by the addition of a measure of actual counseling performance (e.g., rating counseling tapes).

School counselors and counselor educators can use results of this study to take a proactive stance on the issue of the fulfillment of school counselors' professional development needs. The realization that clinical peer consultation and supervision might meet some needs for support and growth could have an empowering effect on the often isolated school counseling practitioner.

References


Author Affiliation

Lori B. Crutchfield is an assistant professor in the Department of Counseling and Clinical Programs at Columbus State University. L. DiAnne Borders is a professor in the Department of Counseling and Educational Development at the University of North Carolina at Greensboro. Edwin Herr served as the action editor for this article. This article is a summary of the doctoral dissertation study of Lori B. Crutchfield, completed at the Department of Counseling and Educational Development of the University of North Carolina at Greensboro under the direction of L. DiAnne Borders. The study was supported in part by a grant from the Small Grants in the Schools program administered through North Carolina State University and was recognized with the 1996 Association for Counselor Education Supervision Outstanding Dissertation Award. The authors thank Jack Culbreth for his assistance in conducting the interventions and the school counselors for their willingness to participate. Correspondence concerning this article should be addressed to Lori B. Crutchfield, Department of Counseling and Clinical Programs, Columbus State University, 4225 University Avenue, Columbus, GA 31907-5645. Electronic mail may be sent via Internet to crutchfield_lori@colstate.edu.