

## The Supervisory Relationship: A Conceptual and Psychometric Review of Measures

By: Jodi L. Tangen, [DiAnne Borders](#)

**This is the peer reviewed version of the following article:**

Tangen, J. L., & Borders, L. D. (2016). The supervisory relationship: A conceptual and psychometric review of measures. *Counselor Education and Supervision*, 55(3), 159-181. doi: 10.1002/ceas.12043,

**which has been published in final form at <http://dx.doi.org/10.1002/ceas.12043>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Self-Archiving.**

**\*\*\*© American Counseling Association. Reprinted with permission. No further reproduction is authorized without written permission from Wiley. This version of the document is not the version of record. Figures and/or pictures may be missing from this format of the document. \*\*\***

### **Abstract:**

To date, a comprehensive review of supervisory relationship measures has yet to be published. In this article, the authors explore conceptualizations of the supervisory relationship, describe and critique 11 measures, provide recommendations for researchers and practitioners when selecting measures, and offer suggestions regarding future measure development.

**Keywords:** clinical supervision | supervisory relationship | supervision measures

### **Article:**

Consistently, the supervisory relationship has been identified as the pivotal component of effective clinical supervision (Borders & Brown, 2005; Goodyear, 2014; Ladany & Muse-Burke, 2001). Researchers (e.g., Beinart, 2014; Lampropoulos, 2002; Weaks, 2002; Worthen & McNeill, 1996; Zarbock, Drews, Bodansky, & Dahme, 2009) have found that a strong supervisory relationship predicts a range of positive outcomes, including increased supervisee disclosure and stronger supervisee–client relationships (Goodyear, 2014). Thus, the essential components of strong and positive supervisory relationships are of great interest to supervision practitioners and researchers, as evidenced by numerous studies cited in Bernard and Goodyear (2014).

The supervisory relationship, however, is a broad, nuanced construct and one that is difficult to describe in full (Borders, 2006; Borders & Brown, 2005). The supervisory relationship incorporates many related constructs, such as relationship development phases, multicultural influences, parallel processes, transference and countertransference issues, as well as idiosyncratic supervisee and supervisor characteristics (Bernard & Goodyear, 2014; Goodyear, 2014; Ladany & Muse-Burke, 2001; Muse-Burke, Ladany, & Deck, 2001). In addition, the impact of supervision-specific factors, such as evaluation, power, and gatekeeping, must be considered. To further complicate matters, the many variables associated with the

supervisory relationship reciprocally influence one another (Ladany & Muse-Burke, 2001). Bernard and Goodyear (2014) perhaps captured the challenge of defining the supervisory relationship best when they stated, “Supervisory relationships are multilayered and complex. To examine them is akin to scanning a forest through a telescope: Each focal range will reveal different aspects and details of the forest” (p. 64).

In conceptualizing the supervision relationship, theorists have attempted to reflect these complexities. The working alliance (Bordin, 1983; Fleming & Benedek, 1964)—consisting of an agreement on goals, an agreement on tasks, and the quality of the supervisor–supervisee bond—seems to be the most readily referenced (Watkins, 2014b). Holloway (1995) focused on the interpersonal relationship, phases of the relationship, and the supervisory contract. Watkins (2011b), drawing from Gelso and Carter's (1994) conceptualization of the therapy relationship, described a tripartite model composed of the alliance, transference–countertransference phenomena, and the real relationship. Such varied conceptualizations certainly reflect the “focal range” (p. 64) noted by Bernard and Goodyear (2014), but are likely confusing for practitioners, educators, and researchers seeking to clarify essential components of the supervisory relationship.

Indeed, given the difficulty in capturing the supervisory relationship conceptually, it is not surprising that it is also difficult to capture empirically. Ellis and Ladany (1997) criticized early attempts to measure the supervisory relationship for heavy reliance on measures of the counseling relationship; these measures essentially substituted the terms *supervisor* and *supervisee* for *counselor* and *client*. This approach highlighted presumed similarities between the supervisory and counseling relationships, but ignored pedagogical and evaluative aspects of the supervision enterprise. More recently, measures based specifically in the supervisory relationship have been published, both in the United States and abroad, reflecting the growing global emphasis on supervision practice and research (Borders et al., 2014). To date, however, the conceptual and psychometric characteristics of these measures have not been examined. Such a review would aid supervision researchers both in choosing valid and reliable instruments and in testing hypotheses of underlying theoretical tenets of measures (Ellis & Ladany, 1997). In addition, a review would provide supervision practitioners with a readily available resource for identifying relationship dynamics to consider and evaluate in their practice.

Accordingly, we offer a systematic and comprehensive review and critique of existing measures of the supervisory relationship, using criteria for rigorous instrument construction based on Ellis, D'Iuso, and Ladany's (2008) guidelines, followed by an overarching evaluation of all of the measures, and then considerations for researchers and practitioners. Although a few older measures were previously reviewed (cf. Ellis & Ladany, 1997), we decided to include them to give readers a fuller scope of available measures.

## **Method**

To identify supervisory relationship measures, we used the research database PsycINFO, with the search words *supervisory relationship* and *supervision relationship*, along with *measure*,

*instrument, scale, and inventory.* We completed a recursive search, first reviewing emergent articles and then using the reference lists of these articles and comprehensive supervision textbooks (e.g., Bernard & Goodyear, 2014) to locate additional measures. Throughout this process, we continually honed our search criteria and eventually settled on four major inclusion criteria. Measures had to be (a) focused on the supervisory relationship, (b) publicly available (i.e., published in peer-reviewed journals, accessible online, or available from the author), (c) written in English, and (d) written for individual (as opposed to group) supervision. With regard to the first criterion, many of the measures we found included attention to the supervisory relationship as part of their focus on other constructs; however, because the relationship did not appear to be the main focus, they were not included in this review (e.g., Supervisory Styles Inventory [Friedlander & Ward, 1984]; Supervisee Attachment Strategies Scale [Menefee, Day, Lopez, & McPherson, 2014]; Collaborative Supervision Behaviors Scale [Rousmaniere & Ellis, 2013]; Psychotherapy Supervisory Inventory [Shanfield, Mohl, Matthews, & Hetherly, 1989]; Multicultural Supervision Competencies Questionnaire [Wong & Wong, as cited in Bernard & Goodyear, 2014]; Questionnaire to Evaluate Supervision [Zarbock et al., 2009]; Relational Behavior Scale [Shaffer & Friedlander, 2015]). Another measure, Woo's (2013) Supervisory Relationship Scale, appeared promising but was written in Chinese. This process yielded 11 measures.

We reviewed the measures using evaluation criteria based on Ellis et al.'s (2008) seven guidelines for best practices in measurement construction (i.e., theorizing, constructs, and supervision context; item pools; content validity data; derivation sample; cross-validity sample; diversity and cross-cultural samples; and further construct validity investigations). To streamline our review, we collapsed their guidelines into three major evaluation criteria for reporting purposes: (a) construct conceptualization and initial measure creation, (b) investigation across sample populations, and (c) validity and reliability statistics. Our critiques are primarily based on the authors' original reports of these measures. We also searched for updated psychometric information published by the first authors of the measures, but did not find additional information. More recent psychometrics (typically only internal consistency) may be found in subsequent studies that use some measures (examples are cited in the sections that follow where applicable). In the following paragraphs, we provide a chronological overview of the measures to reflect the evolution of this work. More detailed information is reported in Table 1.

**Table 1.** Supervisory Relationship Measures

Measure and Sample Item	Purpose and Respondent	Description	Initial Sample	Reliability <sup>a</sup>	Validity Sample <sup>b</sup>
BLRI-S (Schacht et al., 1988) <ul style="list-style-type: none"> <li>“__ respected me” (p. 701).</li> <li>“__ pretended that he liked me or understood me more than he</li> </ul>	<i>Purpose:</i> “Assess the experience of the facilitative conditions in the supervisory	40 items on a 6-point Likert-type scale (M and L forms). <i>Scales:</i> Regard, Unconditionality, Empathic Understanding, Congruence, and	152 members of APA (clinical or counseling psychology, recently received doctorates).	Overall $\alpha = .92$ Scale $\alpha$ : Regard = .85 to .90; Unconditionality = .80 to .82; Empathic Understanding = .75 to .77; Congruence =	Positive correlations ( $r = .64$ to $.95$ ) between the BLRI-S and other therapy versions of the measure. Two forms of the measure tested: M

really did" (p. 701).	relationship" (p. 699). <i>Respondents:</i> Supervisees	Willingness to Be Known		.79 to .83; Willingness to Be Known = .72 to .80	and L. Comparison of the results of the two versions indicated that supervisors used for the M form were more facilitative ( $p < .001$ ) than those used for the L form.
WAI/S (Bahrck, 1990) <ul style="list-style-type: none"> <li>• <i>Supervisor Form:</i> "I believe ___ likes me" (p. 97).</li> <li>• <i>Supervisee Form:</i> "___ and I understand each other" (p. 101).</li> </ul>	<i>Purpose:</i> Measure the strength of the working alliance in supervision as perceived by both supervisors and supervisees. <i>Respondents:</i> Supervisors and supervisees	36 items on a 7-point Likert-type scale. <i>Scales:</i> Goals, Tasks, and Bond	17 supervisees (counseling psychology doctoral students) and 10 supervisors (counseling psychology faculty and advanced graduate students).	None by author. Ladany et al. (1997) ( $n = 105$ ): Overall $\alpha = .93$ Ladany & Friedlander (1995) ( $n = 123$ ): Scale $\alpha$ : Goals = .92; Tasks = .93; Bond = .91	Positive correlation ( $r = .80$ ) between experimental group supervisors' and supervisees' ratings of the working alliance between them. In adapting the instrument, the author had 7 people rate the instrument and found interrater agreement of 97.6% for Bond, 60% for Goals, 64% for Tasks, which the author said suggested two factors (Bond and Goals/Tasks).
SWAI (Efstation et al., 1990) <ul style="list-style-type: none"> <li>• <i>Supervisor Version:</i> "I facilitate my trainee's talking in our sessions" (p. 323).</li> <li>• <i>Trainee Version:</i> "My supervisor helps me talk freely in our sessions" (p. 323).</li> </ul>	<i>Purpose:</i> "Measure the relationship in counselor supervision" (p. 322). <i>Respondents:</i> Supervisors and supervisees	<i>Supervisor Version:</i> 23 items on a 7-point Likert-type scale. <i>Trainee Version:</i> 19 items on a 7-point Likert-type scale. <i>Supervisor scales:</i> Client Focus, Rapport, and Identification <i>Trainee scales:</i> Rapport and Client Focus	185 supervisors and 178 trainees within psychology internship programs.	Supervisor scale $\alpha$ : Client Focus = .71; Rapport = .73; Identification = .77 Trainee scale $\alpha$ : Client Focus = .77; Rapport = .90	Positive correlation ( $r = .50$ and $.52$ , respectively) between the SWAI Client Focus scale and the Task Oriented scale of the Supervisor and Trainee Forms of the Supervisor Styles Inventory (SSI; Friedlander & Ward, 1984). Positive low correlation ( $r = .20$ and $.30$ for

					<p>supervisors and <math>r = .04</math> and <math>.21</math> for trainees, respectively) between the SWAI and the SSI Attractive and Interpersonally Sensitive scales. Low correlation (<math>r = -.06</math> and <math>.00</math>, respectively) between the SWAI Rapport scale and the Task Oriented scale of the Supervisor and Trainee Forms of the SSI. Positive correlation (<math>r = .15</math> and <math>.22</math>, respectively) between the Client Focus and Rapport scales of the Trainee Version of the SWAI and the Self-Efficacy Inventory (Friedlander &amp; Snyder, 1983).</p>
<p>WAI-SR (Smith et al., 2002)</p> <ul style="list-style-type: none"> <li>• “__ and I understand each other.”</li> <li>• “What I am doing in supervision gives me new ways of looking at my counseling sessions.”</li> </ul>	<p><i>Purpose:</i> “Examine how a supervisory relationship develops and matures over time given the three factors of goals, tasks, and bond” (p. 8). <i>Respondents:</i> Supervisees</p>	<p>36 items on a 7-point Likert-type scale. <i>Scales:</i> Goal, Task, and Bond</p>	<p>53 students (child clinical psychology, school psychology, adult clinical psychology, and other programs); over half were master's degree students.</p>	<p>Overall <math>\alpha = .97</math> Scale <math>\alpha</math>: Bond = <math>.91</math>; Goal = <math>.92</math>; Task = <math>.94</math></p>	<p>Positive correlation (<math>r = .76</math> and <math>.73</math>, respectively) between the WAI-SR Bond factor and the SSI Attractive and Interpersonally Sensitive scales. Positive correlation (<math>r = .75</math>) between the WAI-SR Goal factor and the Goal Setting scale of the Evaluation Process Within Supervision</p>

					Inventory (EPSI; Lehrman-Waterman & Ladany, 2001). Positive correlation ( $r = .84$ ) between the WAI-SR Task factor and the Role Ambiguity scale of the Role Conflict and Role Ambiguity Inventory (RCRAI; Olk & Friedlander, 1992). Divergent validity correlations somewhat limited; the WAI-SR Task factor did not correlate with the SSI Task Oriented scale.
<p>FSS (Szymanski, 2003)</p> <ul style="list-style-type: none"> <li>“I am sensitive to the power differences that exist between my supervisees and myself” (p. 225).</li> <li>“I facilitate open, flexible, and egalitarian interactions with my supervisees” (p. 225).</li> </ul>	<p><i>Purpose:</i> “Assess feminist supervision practices in clinical supervision” (p. 221).  <i>Respondents:</i> Supervisors</p>	<p>32 items on a 7-point Likert-type scale.  <i>Scales:</i> Collaborative Relationships, Power Analysis, Emphasis on Diversity and Social Context, and Feminist Advocacy and Activism</p>	<p>Exploratory factor analysis with 108 supervisors; confirmatory factor analysis with 161 supervisors.</p>	<p>Overall <math>\alpha = .95</math>  Scale <math>\alpha</math>: Collaborative Relationships = .72 to .74; Power Analysis = .79 to .85; Emphasis on Diversity and Social Context = .86 to .94; Feminist Advocacy and Activism = .93 to .95</p>	<p>Positive correlation (<math>r = .74</math>) between the FSS and feminist self-identification. Positive correlation (<math>r = .62</math>) between the FSS and the Multicultural Counseling and Awareness Scale (Ponterotto, et al., 2000). Positive correlation (<math>r = .43</math>) between the FSS and the SWAI.</p>
<p>BSAS (Rønnestad &amp; Lundquist, 2009)</p> <ul style="list-style-type: none"> <li><i>Supervisor Form:</i> “I treat my trainee with respect.”</li> <li><i>Trainee Form:</i> “My</li> </ul>	<p><i>Purpose:</i> Brief measure of supervisory alliance.  <i>Respondents:</i> Supervisors</p>	<p><i>Supervisor and Trainee Forms:</i> 12 items on a 6-point Likert-type scale.  <i>Trainee scales:</i> Bond and Coaction</p>	<p>600 Norwegian psychologists (Trainee Form).</p>	<p>Trainee scale <math>\alpha</math>: Bond = .91; Coaction = .93</p>	<p>In process (according to authors).</p>

supervisor treats me with respect.”	and supervisees				
<p>SRSI (Lizzio et al., 2009)</p> <ul style="list-style-type: none"> <li>“Makes an effort to listen and understand even when he/she strongly disagrees with what I am proposing” (p. 134).</li> </ul>	<p>Purpose: Measure the “supervisees’ perceptions of relationship processes” (p. 127).  <i>Respondents:</i> Supervisees</p>	<p>12 items on a 7-point Likert-type scale.  <i>Scales:</i> Support, Challenge, and Openness</p>	<p>261 psychology graduates in Australia.</p>	<p>Scale <math>\alpha</math>: Support = .89; Challenge = .82; Openness = .83</p>	<p>SRSI scores predicted 41.7% of supervisor effectiveness. Openness and Support negatively predicted supervisee anxiety and Challenge positively predicted it.</p>
<p>LASS (Wainwright, 2010)</p> <ul style="list-style-type: none"> <li>“My supervisor and I understood each other in this session” (p. 161).</li> </ul>	<p><i>Purpose:</i> Measure the supervisory alliance from the supervisees’ perspective.  <i>Respondents:</i> Supervisees</p>	<p>3 items on a visual analogue scale.  <i>Item focus:</i> supervision approach, relationship, and helpfulness of the session</p>	<p>Pilot tested on 98 UK clinical psychology students; later validated with 140 UK trainee clinical psychologists.</p>	<p>Overall <math>\alpha</math> = .71  Test-retest <math>r</math> = .63</p>	<p>Positive correlation (<math>r</math> = .71) between the LASS and the SRQ.  Positive correlation (<math>r</math> = .59) between the LASS and the Supervisory Satisfaction Questionnaire (SSQ; Ladany et al., 1996).  Negative correlation (<math>r</math> = -.52) between the LASS and the RCRAI.</p>
<p>SRQ (Palomo et al., 2010)</p> <ul style="list-style-type: none"> <li>“I felt safe in my supervision sessions” (p. 147).</li> <li>“My supervisor appeared interested in me as a person” (p. 148).</li> </ul>	<p><i>Purpose:</i> Measure the supervisory relationship.  <i>Respondents:</i> Supervisees</p>	<p>67 items on 7-point Likert scale.  <i>Scales:</i> Safe Base, Structure, Commitment, Reflective Education, Role Model, and Formative Feedback</p>	<p>284 second- and third-year British doctoral students in psychology. 85 follow-up participants used to examine test-retest reliability.</p>	<p>Overall <math>\alpha</math> = .98  Scale <math>\alpha</math>: Safe Base = .97; Structure = .87; Commitment = .95; Reflective Education = .93; Role Model = .95; Formative Feedback = .93  Test-retest <math>r</math> = .97</p>	<p>Positive correlations (<math>r</math> = .70 to .81) between the SRQ and the EPSI (Lehrman-Waterman &amp; Ladany, 2001).  Negative correlations (<math>r</math> = -.69 to -.64) between the SRQ and the RCRAI.  Positive correlations (<math>r</math> = .86 to .91) between the SRQ and the</p>

					Working Alliance Inventory (WAI; Bahrick, 1990). Positive correlation ( $r = .86$ ) between SRQ and Revised Relationship Inventory (Schacht et al., 1988). The SRQ predicted supervision satisfaction, contribution to personal and professional development, impact on therapy and client progress, and perceived competence of supervisors.
SRM (Pearce et al., 2013) <ul style="list-style-type: none"> <li>“My trainee is open and honest in supervision” (p. 266).</li> <li>“I am able to share my strengths and my weaknesses with my trainee” (p. 268).</li> </ul>	<i>Purpose:</i> Measure supervisors’ perspectives of the supervisory relationship. <i>Respondents:</i> Supervisors	51 items on a 7-point Likert scale. <i>Scales:</i> Safe Base, Supervisor's Professional Commitment to Supervision, Trainee Contribution, External Influences, and Supervisor's Emotional Investment	267 British clinical psychologist supervisors. 134 participants used to examine test-retest reliability.	Overall $\alpha = .90$ Scale $\alpha$ : Safe Base = .96; Supervisor's Professional Commitment to Supervision = .79; Trainee Contribution = .94; External Influences = .71; Supervisor's Emotional Investment = .78 Test-retest $r = .94$	Positive correlations ( $r = .73$ to $.83$ ) between the SRM and the WAI. Positive correlations ( $r = .71$ to $.77$ ) between the SRM and the Trainee Personal Reaction Scale-Revised (Holloway & Wampold, 1984). Positive correlations ( $r = .21$ to $.48$ ) between the SRM and the SSI. Positive correlation ( $r = .85$ and $.68$ , respectively) between the SRM and the authors’ measure of outcome and of

					supervisory satisfaction.
<p>S-SRQ (Cliffe et al., 2014)</p> <ul style="list-style-type: none"> <li>• “My supervisor was approachable.”</li> <li>• “My supervisor was respectful of my views and ideas.”</li> </ul>	<p><i>Purpose:</i> Measure the supervisory relationship (shortened version of the SRQ).</p> <p><i>Respondents:</i> Supervisees</p>	<p>18 items on a 7-point Likert scale.</p> <p><i>Scales:</i> Safe Base, Reflective Education, and Structure</p>	<p>203 clinical psychology trainees in the UK.</p> <p>86 participants used to examine test-retest reliability.</p>	<p>Overall <math>\alpha = .96</math></p> <p>Scale a: Safe Base = .97; Reflective Education = .89; Structure = .88</p> <p>Test-retest <math>r = .94</math></p>	<p>Positive correlation (<math>r = .92</math>) between the S-SRQ and the WAI.</p> <p>Positive correlation (<math>r = .95</math>) between the S-SRQ and the SRQ.</p> <p>Negative correlations (<math>r = -.73</math> to <math>-.68</math>) between the S-SRQ and the RCRAI.</p> <p>No significant correlations (<math>r = -.13</math> to <math>.08</math>) between the S-SRQ and the Short Scale of the Eysenck Personality Questionnaire–Revised (Eysenck et al., 1985).</p> <p>The S-SRQ predicted supervisor satisfaction (based on the SSQ): <math>R^2 = .83</math>.</p> <p>The S-SRQ predicted supervisor effectiveness (based on indices of supervision outcome; Friedlander &amp; Ward, 1984): <math>R^2 = .74</math>.</p>

**Note.** BLRI-S = Barrett-Lennard Relationship Inventory for Supervisory Relationships; APA = American Psychological Association; M form = supervisors who contributed most to supervisees’ therapeutic effectiveness; L form = supervisors who contributed least to supervisees’ effectiveness; WAI/S = Working Alliance Inventory/Supervision; SWAI = Supervisory Working Alliance Inventory; WAI-SR = Working Alliance Inventory of Supervisory Relationships; FSS = Feminist Supervision Scale; BSAS = Brief Supervisory

Alliance Scale; SRSI = Supervisor Relating Style Inventory; LASS = Leeds Alliance in Supervision Scale; UK = United Kingdom; SRQ = Supervisory Relationship Questionnaire; SRM = Supervisory Relationship Measure; S-SRQ = Short Supervisory Relationship Questionnaire. aAlpha is Cronbach's  $\alpha$ . bSee reference list for full statistics.

## **Results**

### **Barrett-Lennard Relationship Inventory for Supervisory Relationships**

Schacht, Howe, and Berman (1988) created the Barrett-Lennard Relationship Inventory for Supervisory Relationships (BLRI-S) to assess supervisees' experiences of facilitative conditions in the supervisory relationship. The BLRI-S was adapted from the BLRI (Barrett-Lennard, 1962), which measures the strength of the therapeutic relationship and is theoretically grounded in Rogers's (1957) core facilitative conditions. The BLRI-S continues to be used by researchers (e.g., supervisor facilitative factors and supervisee personalities [Schacht, Howe, & Berman, 1989], supervisory relationship with substance abuse counselors [Culbreth & Borders, 1999], validity of the Supervisory Relationship Questionnaire [SRQ; Palomo, Beinart, & Cooper, 2010; discussed later], mindfulness in supervision [Daniel, Borders, & Willse, 2015]).

On the basis of our three evaluation criteria, the BLRI-S is somewhat limited. First, in terms of construct conceptualization and initial item creation, the BLRI-S was based on a model of the therapeutic relationship without thorough investigation of the relevance of the core conditions to the supervisory relationship or initial explorations of content validity. Second, Schacht et al. (1988) used a small, developmentally homogeneous sample to initially validate the measure. Third, although some attempt was made to establish construct validity with the BLRI, Schacht et al. (1988) did not explore multiple forms of validity (e.g., convergent validity) or reliability (e.g., test-retest reliability). In addition, some Cronbach's alphas fell below Ellis et al.'s (2008) .80 recommendation, and interscale correlations were high. In fact, because of the latter, Ellis and Ladany (1997) suggested that researchers use only the total score. Despite these limitations, the BLRI-S appears to be one of the strongest in terms of capturing the qualitative essence of the real relationship (Watkins, 2011a, 2012) as distinct from the working aspects of the relationship, and Ellis and Ladany recommended it for practice and research.

### **Working Alliance Inventory/Supervision**

Bahrack (1990) developed the Working Alliance Inventory/Supervision (WAI/S) Supervisor and Supervisee Forms to measure the strength of the supervision working alliance as perceived by both supervisors and supervisees. The measure is theoretically grounded in Bordin's (1983) conceptualization of the supervisory working alliance. Bahrack adapted items directly from Horvath and Greenberg's (as cited in Bahrack, 1990; see also Horvath & Greenberg, 1989) WAI for therapists and clients. The WAI/S continues to be used by researchers (e.g., supervisory alliance and disclosure of countertransference [Mack, 2013], validations of the Multicultural Supervision Inventory [Ortega-Villalobos, 2011], mindfulness in supervision [Daniel et al., 2015]).

According to our evaluation criteria, Bahrnick's (1990) WAI/S is limited in ways similar to the BLRI-S. Both likened the supervisory alliance to the counseling alliance by directly adapting items (from the WAI and BLRI, respectively), thus weakening their construct validity. Second, the small and apparently developmental homogeneity of the initial sample limits external validity. Third, Bahrnick did not report comprehensive reliability and validity statistics; later researchers (Ladany, Brittan-Powell, & Pannu, 1997; Ladany & Friedlander, 1995) reported acceptable levels of internal consistency. Finally, Inman and Ladany (2008) noted multicollinearity of the scales and recommended using only the total score. Some recent researchers (e.g., Crockett & Hays, 2015; Rieck, Callahan, & Watkins, 2015) have used a short form of the WAI/S, a 12-item adapted measure composed of the four highest loading items on each subscale in Tracey and Kokotovic's (1989) factor analysis of the WAI; validity support and acceptable internal consistency for the subscales of this short form have been reported (e.g., Bennett, Mohr, Deal, & Hwang, 2013; Ladany, Mori, & Mehr, 2007), but concerns about construct validity remain.

### **Supervisory Working Alliance Inventory**

Efstation, Patton, and Kardash (1990) created the Supervisory Working Alliance Inventory (SWAI) Supervisor and Trainee Versions to measure the supervisory relationship. The SWAI is grounded in the therapeutic working alliance (see Horvath & Greenberg, 1989) and supervisory working alliance (Bordin, 1983); however, Efstation et al. seemed to be the first to focus more exclusively on the supervisory alliance. They also aimed to capture the social influence present in the supervisory alliance. Efstation et al. recruited assistance from 10 psychology site supervisors to corroborate their initial items and correlated results with other supervision measures. They concluded that more research was needed to validate the scales with other populations. The SWAI has been used in several studies (e.g., wellness in supervision [Storlie & Smith, 2012], alliances in supervision and counseling and trainees' adherence to treatment models [Patton & Kivlighan, 1997], validation of the Supervisor Emphasis Rating Form–Revised [McHenry & Freeman, 1997]).

On the basis of our evaluation criteria, one of the major benefits of the SWAI is Efstation et al.'s (1990) explicit focus on the supervisory working alliance, rather than the therapeutic working alliance. Furthermore, they recruited feedback from supervisors before creating their items. Efstation et al.'s validation sample for the Trainee Version, however, was limited in terms of supervisees' developmental range. Interscale correlations were rather high, subscales accounted for only about one third of the variance, and Cronbach's alphas were low. Accordingly, Ellis and Ladany (1997) did not recommend the SWAI for supervision research or practice; others have suggested using only the composite score (e.g., Patton & Kivlighan, 1997).

### **Working Alliance Inventory of Supervisory Relationships**

In somewhat parallel fashion with Bahrnick's (1990) process, Smith, Younes, and Lichtenberg (2002) created the Working Alliance Inventory of Supervisory Relationships (WAI-SR) on the basis of Bordin's (1983) conceptualization of the supervisory working alliance by slightly revising the wording of the WAI items. However, we could not find follow-up uses of the WAI-

SR. Critiques of this measure are similar to those of the WAI/S, in that the heavy reliance on the therapeutic working alliance makes the measure's construct validity somewhat questionable. Second, as Smith et al. noted, the measure was normed on a very small and developmentally homogeneous sample, thus limiting external validity. In terms of reliability and validity statistics, the measure is promising, especially with regard to concurrent validity and internal consistency.

### **Feminist Supervision Scale**

Szymanski (2003) created the Feminist Supervision Scale (FSS) for supervisors to examine feminist supervision. Szymanski (2003) grounded items in four tenets of feminist supervision: collaborative relationships, analysis of power, diversity and social context, and feminist activism and advocacy. She corroborated these initial items through expert analysis, examined item–total correlations, and conducted an exploratory factor analysis. Szymanski (2003) explored the validity of the FSS in two subsequent studies, one confirming initial convergent validity and the other—a confirmatory factor analysis—further corroborating convergent and discriminant validity. Since its creation, the FSS has been used to explore feminist identity and theories (Szymanski, 2005) and has been modified to explore feminist supervision and self-leadership (Arbel, 2006).

Overall, Szymanski (2003) followed many of Ellis et al.'s (2008) measurement construction guidelines. She carefully conceptualized, theoretically grounded, defined, and content validated the construct of interest; performed two studies with different sample populations for cross-validation; and explored reliability and multiple forms of validity. Still, there are some limitations. First, the measure is grounded in tenets of feminism, which emphasizes an egalitarian relationship. In fact, the five questions that compose the Collaborative Relationships subscale all address issues of power and equality. The power differential in the supervisory relationship is undeniable and needs to be addressed (Falender, 2010); however, it is certainly not the only component of a supervisory relationship. Furthermore, sometimes supervisees (especially beginners) need more hierarchical interventions (Prouty, Thomas, Johnson, & Long, 2001) and directive approaches (Borders & Brown, 2005). Thus, the FSS may not be sensitive to developmental shifts in the supervisory relationship. Second, although Szymanski (2003) used two samples, both were limited in size and diversity. Finally, as Szymanski (2003) noted, the measure would benefit from more exploration of its validity (e.g., predictive validity) and reliability (e.g., test–retest reliability). Nevertheless, the FSS appears to be a reliable and valid measure of feminist supervision.

### **Brief Supervisory Alliance Scale**

Rønnestad and Lundquist (2009) created the Brief Supervisory Alliance Scale (BSAS) Supervisor and Trainee Forms as a brief measure of the supervisory alliance. Although theoretical and psychometric information has yet to be published by the authors, the measure is recommended in Wheeler, Aveline, and Barkham's (2011) common tool kit of practice-based supervision research. Rønnestad and Lundquist explored the internal consistency of the Trainee Form using Norwegian psychologists, which appeared promising; however, information about the reliability of the Supervisor Form and validity for either form is currently lacking.

Furthermore, we could not find other studies using the measure. In short, the BSAS shows promise, but much more information is needed.

### **Supervisor Relating Style Inventory**

Lizzio, Wilson, and Que (2009) created the Supervisor Relating Style Inventory (SRSI) to measure supervisees' perceptions of the supervisory relationship across disciplines (e.g., counseling, teaching, nursing). Lizzio et al. chose three broad relational elements—challenge, support, and openness—as the basis of their inventory. They wrote 21 initial items based on their knowledge of the relationship elements, sought feedback from eight supervisors and supervisees, tested the measure with Australian psychology graduates, conducted factor analyses, and reduced their items to 12 on the basis of factor loadings. However, we were unable to find subsequent uses of the measure.

According to our evaluation criteria, the SRSI yields mixed results. Although Lizzio et al. (2009) defined their constructs, they apparently did not ground them in a theoretical framework. Furthermore, much of the literature they reviewed to create the measure was based in psychology; they included only psychology graduate students in their validation study (homogeneity) and did not cross-validate across settings, which somewhat contradicts their intention of creating an interdisciplinary measure. Still, a strength of the SRSI is that Lizzio et al. created it specifically for the supervisory relationship. Lizzio et al.'s investigations of predictive validity and internal consistency are relatively sound; however, explorations of convergent and discriminant validity and test–retest reliability of this promising measure are needed.

### **Leeds Alliance in Supervision Scale**

Wainwright (2010) created the Leeds Alliance in Supervision Scale (LASS) as a very brief measure that could be used by supervisees after every supervision session. He drew from four theories of the supervisory alliance: Bordin's (1983) conceptualization of the supervisory working alliance, Holloway's (1997) systems approach, Beinar's (2002) grounded theory study of the supervisory relationship, and Palomo et al.'s (2010) SRQ. Wainwright collected items from existing supervisory measures, had coders group them into themes, and then rated items on the basis of how well they represented the themes. Wainwright then modified the items, pilot tested them with clinical psychology trainees in the United Kingdom, and conducted a principal components analysis and cluster analysis that yielded three major clusters. Wainwright explored psychometrics of the measure by comparing his results with those for associated measures and by exploring internal consistency and test–retest reliability; he concluded that the measure displayed adequate reliability and validity. Since its creation, the LASS has been used to explore the relationship between racially matched and nonmatched supervisors and supervisees (Payne, Smith, Tuchfeld, & Suprina, 2013) and has been recommended to explore feedback in supervision (Redfern, 2014).

On the basis of our evaluation criteria, we consider the LASS a promising measure. Wainwright (2010) carefully conceptualized the construct based on four theories and items from similar instruments. He tested the measure on two different samples and explored multiple forms of validity (i.e., concurrent, convergent, and discriminant validity) and reliability (i.e., internal

consistency and test–retest reliability). Still, the LASS is limited in a few ways. First, with only three items, it likely does not capture the breadth and depth of the supervisory relationship. Furthermore, the relationship appears to be associated with only one of the items, suggesting that Wainwright conceptualized the relationship as a subset of the supervisory working alliance—contrary to other views (Watkins, 2014b). Second, although Wainwright used two samples, they were both limited in size and diversity. Finally, as he acknowledged, the Cronbach's alpha and test–retest statistics are low, and further investigations of construct and predictive validity would be beneficial. If these limitations were addressed, the LASS could be a very viable measure, especially for supervision practitioners.

### **SRQ and Short SRQ**

Palomo et al. (2010) aimed to create a measure of the supervisory relationship from supervisees' perspectives. The SRQ is theoretically based in an earlier grounded theory study conducted by Beinart (as cited in Palomo et al., 2010; see also Beinart, 2002), who examined supervisees' descriptions of supervisor characteristics that affected their therapeutic effectiveness. Palomo et al. wrote 111 items based on Beinart's (2002) results. To test their measure, they sent it along with other supervision measures and their own indices of supervision outcome to a sample of 2nd- and 3rd-year British doctoral students in clinical psychology. Palomo et al. then conducted a principal components analysis and extracted six factors. Although we were unable to find subsequent uses of the measure, Watkins (2014a) and Lewis, Scott, and Hendricks (2014) mentioned the SRQ as a viable measure.

For the most part, Palomo et al. (2010) met many of the evaluation criteria. They conceptualized the construct and created items based on a specific focus on a model of the supervisory relationship, and they used a large (over 200) sample. Finally, they investigated multiple types of validity (i.e., convergent, discriminant, and predictive validity) and reliability (i.e., internal consistency and test–retest reliability); results indicated that the measure is psychometrically sound. Palomo et al. also acknowledged a few limitations of the measure—the main one being the developmentally homogeneous nature of the sample, which could limit external validity.

Another possible limitation of the SRQ is its length (67 items). To address this concern, Cliffe, Beinart, and Cooper (2014) developed the Short SRQ (S-SRQ), an 18-item version of the SRQ. These authors reduced the number of items by examining external and internal item quality and obtaining feedback from an experienced supervision researcher. They then administered the initial draft and several other measures to trainee clinical psychologists in the United Kingdom. An exploratory principal components analysis yielded three factors. Cliffe et al. also explored other psychometrics (i.e., internal consistency; convergent, divergent, and predictive validity; and test–retest reliability) and reported sound validity and reliability. However, we were unable to find subsequent uses of the S-SRQ.

Strengths and limitations of the S-SRQ parallel those of the SRQ. The S-SRQ was created based on a model of the supervisory relationship, and Cliffe et al. (2014) used a large sample to validate the measure; however, as with the SRQ, this sample was developmentally homogeneous and Cliffe et al. did not cross-validate with a new sample. However, like the SRQ, the measure

appears to have strong validity (i.e., convergent, divergent, and predictive validity) and reliability (i.e., internal consistency and test–retest reliability). With further investigations with more diverse samples, the S-SRQ appears to have much potential.

### **Supervisory Relationship Measure**

Pearce, Beinart, Clohessy, and Cooper (2013) created the Supervisory Relationship Measure (SRM) to measure supervisors' perspectives of the supervisory relationship. They used three core categories from Clohessy's (2008) grounded theory study of 12 clinical psychologist supervisors to create items: core relational factors, flow of supervision, and contextual influences. Pearce et al. then examined face validity, pilot tested the measure with volunteers, and sent it along with four other questionnaires to British clinical psychology supervisors. On the basis of a principal components analysis, a factor analysis, and item loadings, the authors retained five factors and then examined multiple forms of validity and reliability. Although we could not find subsequent uses of the measure, it has been mentioned in recent discussions of the supervisory relationship (Falendar & Shafranske, 2014; Watkins, 2014a).

The SRM appears to be a promising measure in light of our evaluation criteria. Pearce et al. (2013) created the measure based solely on conceptualizations of the supervisory relationship and used a large sample (more than 200) to initially test it. They explored multiple forms of validity (i.e., convergent, divergent, predictive, and concurrent validity) and reliability (i.e., internal consistency and test–retest reliability); results indicated that it was a sound measure of the supervisory relationship. Nevertheless, Pearce et al. acknowledged limitations, including item creation procedures (i.e., based on one qualitative study), a lack of diversity in the validation sample, reliance on a self-created measure of outcome and satisfaction to establish validity, and issues with some of the statistical procedures. Another possible limitation is the length (51 items); however, with an exclusive focus on the supervisory relationship and fairly sound psychometrics, the SRM could be of benefit to researchers and practitioners.

### **Discussion**

Our review of the 11 measures certainly illustrated the varied “focal range” (Bernard & Goodyear, 2014, p. 64) of the supervisory relationship. To synthesize and extend our critiques of each measure, we provide (a) an overarching evaluation based on our three criteria from Ellis et al.'s (2008) measurement construction guidelines, followed by (b) instrument selection considerations for researchers and practitioners.

### **Overarching Evaluation**

First, in terms of construct conceptualization and initial measure creation, the measures generally improved over time. Authors of earlier measures (e.g., BLRI-S, WAI/S) adapted or wrote items based on conceptualizations of the therapeutic relationship—a questionable approach—whereas authors of more recent measures (e.g., SRQ, S-SRQ, SRM) focused exclusively on conceptualizations of the supervisory relationship. Many authors also chose a theoretical framework to ground their measure (most using Bordin's, 1983, conceptualization of the supervisory working alliance). Although these are strengths, we sometimes found the operational

definitions and clear boundary demarcations of the supervisory relationship lacking. This finding is also corroborated in other studies (Kemer, Borders, & Willse, 2014; Olds & Hawkins, 2014), in which the supervisory relationship appeared to pervade many other supervision components. Future researchers seeking to create new measures need to clarify the elements of exactly what is being measured. Finally, in most cases, we noted that fewer than the recommended minimum 30 experts (Ellis et al., 2008) were used to explore the content validity of the initial items.

In addition, several variables that characterize the supervisory relationship (see Ladany & Muse-Burke, 2001; Muse-Burke et al., 2001) were lacking in the measures. Only the FSS addressed multicultural issues, power was directly addressed only in the FSS and SRQ, and direct questions about transference and countertransference were not found in any of the measures. We also found minimal attention to the depth of the supervisory relationship as characterized by Watkins's (2011a, 2012) descriptions of the real relationship. Although later measures (e.g., SRQ, SRM) certainly included relational elements, they nevertheless did not seem to capture the potential transformative power of the supervisory relationship (cf. Ladany et al., 2012). In addition, many (but not all) of the measures provide only a static view of the relationship, ignoring the ongoing negotiation of the relationship (cf. Doran, Safran, Waizmann, Bolger, & Muran, 2012), as well as supervisor responsiveness to supervisee needs (Friedlander, 2012) and conflicts in the relationship (Ladany, Friedlander, & Nelson, 2005).

Regarding our second evaluation criterion, sampling procedures lacked rigor across the measures. Samples sizes in many initial validation studies were below 200, and many of the authors did not readily cross-validate the measures. In addition, and perhaps most important, the authors typically used homogeneous samples, especially with regard to supervisees' developmental level and diversity.

Finally, Ellis et al. (2008) highlighted the importance of thorough validity and reliability evaluation. Overall, the psychometrics of the measures not only have improved over time, but also have become more rigorous and robust. For example, investigations of multiple types of validity were lacking for the BLRI-S and WAI/S, whereas for later measures (e.g., SRQ, S-SRQ, SRM) multiple forms (especially construct and criterion validity) were examined. Similarly, investigations of test-retest reliability were lacking for earlier measures (e.g., BLRI-S, WAI/S, SWAI, FSS, SRSI) but were apparent with more recent ones (e.g., LASS, SRQ, S-SRQ, SRM). The reported internal consistency of most of the measures was rather sound (Cronbach's alphas above .80), although the Cronbach's alphas for the SWAI; the LASS; and a few subscales on the BLRI-S, FSS, and SRM were below .80. At this point, the SRQ, S-SRQ, and SRM seem especially exhaustive in investigations of validity and reliability and thus may be the most viable choices for researchers and practitioners with regard to construction criteria.

### **Instrument Selection Considerations**

In light of our evaluation results, it seems prudent that researchers and practitioners be intentional in choosing a measure. For empirical work, researchers could (a) determine their purpose of measurement and the specific elements of the relationship they desire to measure (e.g., if wanting an instrument that measures some aspect of power, choose the FSS or SRQ; if

wanting a broad check-in for use across multiple sessions, choose the LASS; if wanting an educational perspective, choose the SRQ), (b) determine for whom the measure is intended (e.g., choose the SRSI, LASS, SRQ, or S-SRQ for supervisees; choose the FSS or SRM for supervisors), (c) consider psychometrics of the measure (generally choose more recent instruments [e.g., SRQ, S-SRQ, SRM], which are based on more robust construction designs), (d) consider the length of the measure (e.g., the SRM [51 items] and the SRQ [67 items] may be too long, whereas the LASS [three items] may be too short), (e) closely examine the appropriateness of the items (e.g., the Trainee Contribution subscale of the SRM with items such as “My trainee is able to hold an appropriate caseload” [Pearce et al., 2013, p. 267] may not be applicable), and (f) make an informed decision.

Similarly, supervision practitioners can evaluate which measures might be helpful in initiating an upfront conversation about the supervisory relationship, such as what relationship dimensions are desired, how the dyad can work toward that goal, and ways they will communicate what is and is not working in the relationship. In line with recommendations regarding regular use of session outcome measures in clinical work, supervisors may also invite ongoing feedback about the relationship via one of the shorter measures (e.g., LASS).

### **Limitations**

We acknowledge limitations in our review. First, because a distinct definition of the supervisory relationship (and its relation to the supervisory working alliance) is somewhat unclear, it was difficult to establish definite inclusion criteria for our measures; thus, other researchers may have selected different measures. Second, we established criteria that may have prematurely excluded measures that could not be located online or obtained from authors. Third, our English-only measures may not address important relationship dynamics in non-English cultures. Finally, because of space limitations, we could not provide exhaustive summaries, critiques, and updated psychometric information for each measure.

### **Conclusion**

The supervisory relationship is the pivotal component of supervision (Borders & Brown, 2005; Goodyear, 2014; Ladany & Muse-Burke, 2001), and selecting a measure of it for whatever purpose involves multiple considerations. We have endeavored to outline some of these considerations and provide a resource for measure selection. We commend authors’ efforts to improve measures of the supervisory relationship and hope that this review encourages further advances in measure construction and validation.

At the same time, we continue to question, as have other researchers (e.g., Borders, 2006; Olds & Hawkins, 2014; Watkins, 2011a), whether supervision scholars have yet achieved a comprehensive depiction of the breadth and depth, the complexity and simplicity, of the supervisory relationship. It may be that, in creating a supervision measure, researchers must choose a “focal range” (Bernard & Goodyear, 2014, p. 64) that reveals limited details of the forest; nevertheless, a broader perspective of the supervisory relationship forest is warranted as well.

## References

\*References marked with an asterisk indicate measures reviewed.

Arbel, O. (2006). Supervisees' perceptions of feminist supervision, self-leadership and satisfaction with supervision: An exploratory study (Doctoral dissertation). Available from ProQuest Dissertations and Theses database.

\*Bahrlick, A. S. (1990). Role induction for counselor trainees: Effects on the supervisory working alliance (Doctoral dissertation). Retrieved from <https://etd.ohiolink.edu/>

Barrett-Lennard, G. T. (1962). Dimensions of therapist response as causal factors in therapeutic change. *Psychological Monographs: General and Applied*, **76**, 1–36. doi:10.1037/h0093918

Beinart, H. (2002). *An exploration of the factors which predict the quality of the relationship in clinical supervision (Unpublished doctoral thesis)*. Open University.

Beinart, H. (2014). Building and sustaining the supervisory relationship. In C. E. Watkins Jr. & D. L. Milne (Eds.), *The Wiley international handbook of clinical supervision* (pp. 257–281). doi:10.1002/9781118846360.ch11

Bennett, S., Mohr, J., Deal, K. H., & Hwang, J. (2013). Supervisor attachment, supervisory working alliance, and affect in social work field instruction. *Research on Social Work Practice*, **23**, 199–209. doi:10.1177/1049731512468492

Bernard, J. M., & Goodyear, R. K. (2014). *Fundamentals of clinical supervision* (5th ed.). Boston, MA: Pearson.

Borders, L. D. (2006). Snapshot of clinical supervision in counseling and counselor education: A five-year review. *The Clinical Supervisor*, **24**, 69–113. doi:10.1300/J001v24n01\_05

Borders, L. D., & Brown, L. L. (2005). *The new handbook of counseling supervision*. Mahwah, NJ: Erlbaum.

Borders, L. D., Glossoff, H. L., Welfare, L. E., Hays, D. G., DeKruyf, L., Fernando, D. M., & Page, B. (2014). Best practices in clinical supervision: Evolution of a counseling specialty. *The Clinical Supervisor*, **33**, 26–44. doi:10.1080/07325223.2014.905225

Bordin, E. S. (1983). A working alliance based model of supervision. *The Counseling Psychologist*, **11**, 35–42. doi:10.1177/0011000083111007

\*Cliffe, T., Beinart, H., & Cooper, M. (2014). Development and validation of a short version of the Supervisory Relationship Questionnaire. *Clinical Psychology and Psychotherapy*. Advance online publication. doi:10.1002/cpp.1935

Clohesy, S. (2008). Supervisors' perspectives on their supervisory relationships: A qualitative analysis (Doctoral thesis, University of Hull, Kingston upon Hull, England). Retrieved from <http://core.kmi.open.ac.uk/display/2731399>

- Crockett, S., & Hays, D. G. (2015). The influence of supervisor multicultural competence on the supervisory working alliance, supervisee counseling self-efficacy, and supervisee satisfaction with supervision: A mediation model. *Counselor Education and Supervision, 54*, 258–273. doi:10.1002/ceas.12025
- Culbreth, J. R., & Borders, L. D. (1999). Perceptions of the supervisory relationship: Recovering and nonrecovering substance abuse counselors. *Journal of Counseling & Development, 77*, 330–338. doi:10.1002/j.1556-6676.1999.tb02456.x
- Daniel, L., Borders, L. D., & Willse, J. (2015). The role of supervisors' and supervisees' mindfulness in clinical supervision. *Counselor Education and Supervision, 54*, 221–232. doi:10.1002/ceas.12015
- Doran, J. M., Safran, J. D., Waizmann, V., Bolger, K., & Muran, J. C. (2012). The Alliance Negotiation Scale: Psychometric construction and preliminary reliability and validity analysis. *Psychotherapy Research, 22*, 710–719. doi:10.1080/10503307.2012.709326
- \*Efstation, J. F., Patton, M. J., & Kardash, C. M. (1990). Measuring the working alliance in counselor supervision. *Journal of Counseling Psychology, 37*, 322–329. doi:10.1037/0022-0167.37.3.322
- Ellis, M. V., D'Iuso, N., & Ladany, N. (2008). State of the art in the assessment, measurement, and evaluation of clinical supervision. In A. K. Hess, K. D. Hess, & T. H. Hess (Eds.), *Psychotherapy supervision: Theory, research, and practice* (pp. 473–499). Hoboken, NJ: Wiley.
- Ellis, M. V., & Ladany, N. (1997). Inferences concerning supervisees and clients in clinical supervision: An integrative review. In C. E. Watkins Jr. (Ed.), *Handbook of psychotherapy supervision* (pp. 447–507). New York, NY: Wiley.
- Eysenck, S. B. G., Eysenck, H. J., & Barrett, P. (1985). A revised version of the Psychoticism Scale. *Personality and Individual Differences, 6*, 21–29. doi:10.1016/0191-8869(85)90026-1
- Falender, C. A. (2010). Relationship and accountability: Tensions in feminist supervision. *Women & Therapy, 33*, 22–41. doi:10.1080/02703140903404697
- Falender, C. A., & Shafranske, E. P. (2014). Clinical supervision: The state of the art. *Journal of Clinical Psychology, 70*, 1030–1041. doi:10.1002/jclp.22124
- Fleming, J., & Benedek, T. (1964). Supervision: A method of teaching psychoanalysis. Preliminary report. *The Psychoanalytic Quarterly, 33*, 71–96.
- Friedlander, M. L. (2012). Therapist responsiveness: Mirrored in supervisor responsiveness. *The Clinical Supervisor, 31*, 103–119. doi:10.1080/07325223.2012.675199
- Friedlander, M. L., & Snyder, J. (1983). Trainees' expectations for the supervisory process: Testing a developmental model. *Counselor Education and Supervision, 22*, 342–348. doi:10.1002/j.1556-6978.1983.tb01771.x

- Friedlander, M. L., & Ward, L. G. (1984). Development and validation of the Supervisory Styles Inventory. *Journal of Counseling Psychology*, **31**, 541–557. doi:10.1037/0022-0167.31.4.541
- Gelso, C. J., & Carter, J. A. (1994). Components of the psychotherapy relationship: Their interaction and unfolding during treatment. *Journal of Counseling Psychology*, **41**, 296–306. doi:10.1037/0022-0167.41.3.296
- Goodyear, R. G. (2014). Supervision as pedagogy: Attending to its essential instructional and learning processes. *The Clinical Supervisor*, **33**, 82–99. doi:10.1080/07325223.2014.918914
- Holloway, E. L. (1995). *Clinical supervision: A systems approach*. Thousand Oaks, CA: Sage.
- Holloway, E. L. (1997). Structures for the analysis and teaching of supervision. In C. E. Watkins Jr. (Ed.), *Handbook of psychotherapy supervision* (pp. 249–276). New York, NY: Wiley.
- Holloway, E. L., & Wampold, B. E. (1984, August). Dimensions of satisfaction in the supervision interview. Paper presented at the meeting of the American Psychological Association, Toronto, Ontario, Canada.
- Horvath, A. O., & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology*, **36**, 223–233. doi:10.1037/0022-0167.36.2.223
- Inman, A. G., & Ladany, N. (2008). Research: The state of the field. In A. K. Hess, K. D. Hess, & T. H. Hess (Eds.), *Psychotherapy supervision: Theory, research, and practice* (pp. 500–517). Hoboken, NJ: Wiley.
- Kemer, G., Borders, L. D., & Willse, J. (2014). Cognitions of expert supervisors in academe: A concept mapping approach. *Counselor Education and Supervision*, **53**, 2–18. doi:10.1002/j.1556-6978.2014.00045.x
- Ladany, N., Brittan-Powell, C. S., & Pannu, R. K. (1997). The influence of supervisory racial identity interaction and racial matching on the supervisory working alliance and supervisee multicultural competence. *Counselor Education and Supervision*, **36**, 284–304. doi:10.1002/j.1556-6978.1997.tb00396.x
- Ladany, N., & Friedlander, M. L. (1995). The relationship between the supervisory working alliance and trainees' experience of role conflict and role ambiguity. *Counselor Education and Supervision*, **34**, 220–231. doi:10.1002/j.1556-6978.1995.tb00244.x
- Ladany, N., Friedlander, M. L., & Nelson, M. L. (2005). *Critical events in psychotherapy supervision: An interpersonal approach*. Washington, DC: American Psychological Association.
- Ladany, N., Hill, C. E., Corbett, M. M., & Nutt, E. A. (1996). Nature, extent, and importance of what psychotherapy trainees do not disclose to their supervisors. *Journal of Counseling Psychology*, **43**, 10–24. doi:10.1037/0022-0167.43.1.10
- Ladany, N., Inman, A. G., Hill, C. E., Know, S., Crook-Lyon, R. E., Thompson, B. J., ... Walker, J. A. (2012). Corrective relational experiences in supervision. In L. G. Castonguay & C. E. Hill (Eds.), *Transformation in psychotherapy: Corrective experiences*

across cognitive behavioral, humanistic, and psychodynamic approaches (pp. 335–352). Washington, DC: American Psychological Association.

Ladany, N., Mori, Y., & Mehr, K. E. (2007). *Trainee perceptions of effective and ineffective supervision interventions*. Unpublished manuscript, Lehigh University, Bethlehem, PA.

Ladany, N., & Muse-Burke, J. L. (2001). Understanding and conducting supervision research. In L. J. Bradley & N. Ladany (Eds.), *Counselor supervision: Principles, process, and practice* (3rd ed., pp. 304–329). Philadelphia, PA: Brunner-Routledge.

Lampropoulos, G. K. (2002). A common factors view of counseling supervision process. *The Clinical Supervisor*, **21**, 77–95. doi:10.1300/J001v21n01\_06

Lehrman-Waterman, D., & Ladany, N. (2001). Development and validation of the Evaluation Process Within Supervision Inventory. *Journal of Counseling Psychology*, **48**, 168–177. doi:10.1037/0022-0167.48.2.168

Lewis, C. C., Scott, K. E., & Hendricks, K. E. (2014). A model and guide for evaluating supervision outcomes in cognitive-behavioral therapy-focused training programs. *Training and Education in Professional Psychology*, **8**, 165–173. doi:10.1037/tep0000029

\*Lizzio, A., Wilson, K., & Que, J. (2009). Relationship dimensions in the professional supervision of psychology graduates: Supervisee perceptions of processes and outcome. *Studies in Continuing Education*, **31**, 127–140. doi:10.1080/01580370902927451

Mack, S. (2013). Supervisory alliance and countertransference disclosure in peer supervision (Doctoral dissertation). Retrieved from <http://pepperdine.contentdm.oclc.org/cdm/ref/collection/p15093coll2/id/244>

McHenry, S., & Freeman, B. (1997). A multimethod-multitrait validation study of the Supervisor Emphasis Rating Form-Revised. *The Clinical Supervisor*, **15**, 1–17. doi:10.1300/J001v15n02\_01

Menefee, D. S., Day, S. X., Lopez, F. G., & McPherson, R. H. (2014). Preliminary development and validation of the Supervisee Attachment Strategies Scale (SASS). *Journal of Counseling Psychology*, **61**, 232–240. doi:10.1037/a0035600

Muse-Burke, J. L., Ladany, N., & Deck, M. D. (2001). The supervisory relationship. In L. J. Bradley & N. Ladany (Eds.), *Counseling supervision: Principles, process, and practice* (3rd ed., pp. 28–62). Philadelphia, PA: Brunner-Routledge.

Olds, K., & Hawkins, R. (2014). Precursors to measuring outcomes in clinical supervision: A thematic analysis. *Training and Education in Professional Psychology*, **8**, 158–164. doi:10.1037/tep0000034

Olk, M. E., & Friedlander, M. L. (1992). Trainees' experiences of role conflict and role ambiguity in supervisory relationships. *Journal of Counseling Psychology*, **39**, 389–397. doi:10.1037/0022-0167.39.3.389

- Ortega-Villalobos, L. (2011). Confirmation of the structure and validity of the Multicultural Supervision Inventory. *Dissertation Abstracts International: Section B. Sciences and Engineering*, **72**(3), 1802.
- \*Palomo, M., Beinart, H., & Cooper, M. J. (2010). Development and validation of the Supervisory Relationship Questionnaire (SRQ) in UK trainee clinical psychologists. *British Journal of Clinical Psychology*, **49**, 131–149. doi:10.1348/014466509X441033
- Patton, M. J., & Kivlighan, D. R. (1997). Relevance of the supervisory alliance to the counseling alliance and to treatment adherence in counselor training. *Journal of Counseling Psychology*, **44**, 108–115. doi:10.1037/0022-0167.44.1.108
- Payne, J. H., Smith, S. D., Tuchfeld, B., & Suprina, J. S. (2013). Relationship development between racially matched and non-matched counselor supervisors and practicum supervisees: Preliminary findings. *Practitioner Scholar: Journal of Counseling & Professional Psychology*, **2**, 108–118. Retrieved from <http://www.thepractitionerscholar.com/index>
- \*Pearce, N., Beinart, H., Clohessy, S., & Cooper, M. (2013). Development and validation of the Supervisory Relationship Measure: A self-report questionnaire for use with supervisors. *British Journal of Clinical Psychology*, **52**, 249–268. doi:10.1111/bjc.12012
- Ponterotto, J. G., Gretchen, D., Utsey, S. O., Reiger, B. P., & Autsin, R. (2000). A construct validity study of the Multicultural Counseling Awareness Scale (MCAS). Unpublished manuscript.
- Prouty, A. M., Thomas, V., Johnson, S., & Long, J. K. (2001). Methods of feminist family therapy supervision. *Journal of Marital and Family Therapy*, **27**, 85–97. doi:10.1111/j.1752-0606.2001.tb01141.x
- Redfern, E. (2014). Feedback in supervision. *Therapy Today*, **25**, 26–29.
- Rieck, T., Callahan, J. L., & Watkins, C. E., Jr. (2015). Clinical supervision: An exploration of possible mechanisms of action. *Training and Education in Professional Psychology*, **9**, 187–194. doi:10.1037/tep0000080
- Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology*, **21**, 95–103. doi:10.1037/h0045357
- \*Rønnestad, M. H., & Lundquist, K. (2009). *The Brief Supervisory Alliance Scale*. Unpublished manuscript, Department of Psychology, Oslo, Norway.
- Rousmaniere, T. G., & Ellis, M. V. (2013). Developing the construct and measure of collaborative clinical supervision: The supervisee's perspective. *Training and Education in Professional Psychology*, **7**, 300–308. doi:10.1037/a0033796
- \*Schacht, A. J., Howe, H. E., & Berman, J. J. (1988). A short form of the Barrett-Lennard Relationship Inventory for Supervisory Relationships. *Psychological Reports*, **63**, 699–706. doi:10.2466/pr0.1988.63.3.699

Schacht, A. J., Howe, H. E., & Berman, J. J. (1989). Supervisor facilitative conditions and effectiveness as perceived by thinking- and feeling-type supervisees. *Psychotherapy: Theory, Research, Practice, Training*, **26**, 475–483. doi:10.1037/h0085466

Shaffer, K. S., & Friedlander, M. L. (2015). What do “interpersonally sensitive” supervisors do and how do supervisees experience a relational approach to supervision? *Psychotherapy Research*. Advance online publication. doi:10.1080/10503307.2015.1080877

Shanfield, S. B., Mohl, P. C., Matthews, K., & Hetherly, V. (1989). A reliability assessment of the Psychotherapy Supervisory Inventory. *The American Journal of Psychiatry*, **146**, 1447–1450. doi:10.1176/ajp.146.11.1447

\*Smith, T. R., Younes, L. K., & Lichtenberg, J. W. (2002, August). Examining the working alliance in supervisory relationships: The development of the Working Alliance Inventory of Supervisory Relationships. Paper presented at the meeting of the American Psychological Association, Chicago, IL.

Storlie, C. A., & Smith, C. (2012). The effects of a wellness intervention in supervision. *The Clinical Supervisor*, **31**, 228–239. doi:10.1080/07325223.2013.732504

\*Szymanski, D. M. (2003). The Feminist Supervision Scale: A rational/theoretical approach. *Psychology of Women Quarterly*, **27**, 221–232. doi:10.1111/1471-6402.00102

Szymanski, D. M. (2005). Feminist identity and theories as correlates of feminist supervision practices. *The Counseling Psychologist*, **33**, 729–747. doi:10.1177/0011000005278408

Tracey, T. J., & Kokotovic, A. M. (1989). Factor structure of the Working Alliance Inventory. *Psychological Assessment: A Journal of Consulting and Clinical Psychology*, **1**, 207–210. doi:10.1037/1040-3590.1.3.207

\*Wainwright, N. A. (2010). *The development of the Leeds Alliance in Supervision Scale (LASS): A brief sessional measure of the supervisory alliance* (Doctoral thesis, University of Leeds, Leeds, England). Retrieved from <https://core.ac.uk/download/files/139/1145290.pdf>

Watkins, C. E., Jr. (2011a). The real relationship in psychotherapy supervision. *American Journal of Psychotherapy*, **65**, 99–116.

Watkins, C. E., Jr. (2011b). Toward a tripartite vision of supervision for psychoanalysis and psychoanalytic psychotherapies: Alliance, transference-countertransference configuration, and real relationship. *Psychoanalytic Review*, **98**, 557–590. doi:10.1521/prev.2011.98.4.557

Watkins, C. E., Jr. (2012). Moments of real relationship in psychoanalytic supervision. *The American Journal of Psychoanalysis*, **72**, 251–268. doi:10.1057/ajp.2012.14

Watkins, C. E., Jr. (2014a). Clinical supervision in the 21st century: Revisiting pressing needs and impressing possibilities. *American Journal of Psychotherapy*, **68**, 251–272.

- Watkins, C. J., Jr. (2014b). The supervisory alliance as quintessential integrative variable. *Journal of Contemporary Psychotherapy*, **44**, 151–161. doi:10.1007/s10879-013-9252-x
- Weaks, D. (2002). Unlocking the secrets of ‘good supervision’: A phenomenological exploration of experienced counsellors’ perceptions of good supervision. *Counselling & Psychotherapy Research*, **2**, 33–39. doi:10.1080/14733140212331384968
- Wheeler, S., Aveline, M., & Barkham, M. (2011). Practice-based supervision research: A network of researchers using a common toolkit. *Counselling & Psychotherapy Research*, **11**, 88–96. doi:10.1080/14733145.2011.562982
- Woo, S. (2013). Development and validation of a Supervisory Relationship Scale. *Chinese Journal of Guidance and Counseling*, **38**, 117–148.
- Worthen, V., & McNeill, B. W. (1996). A phenomenological investigation of “good” supervision events. *Journal of Counseling Psychology*, **43**, 25–34. doi:10.1037/0022-0167.43.1.25
- Zarbock, G., Drews, M., Bodansky, A., & Dahme, B. (2009). The evaluation of supervision: Construction of brief questionnaires for the supervisor and the supervisee. *Psychotherapy Research*, **19**, 194–204. doi:10.1080/10503300802688478