Development of the Chae Optimal Supervision Environment Test

Ki Byung Chae
William & Mary - School of Education

Follow this and additional works at: https://scholarworks.wm.edu/etd
Part of the Educational Leadership Commons, Quantitative Psychology Commons, and the Student Counseling and Personnel Services Commons

Recommended Citation
https://scholarworks.wm.edu/etd/1550154036

This Dissertation is brought to you for free and open access by the Theses, Dissertations, & Master Projects at W&M ScholarWorks. It has been accepted for inclusion in Dissertations, Theses, and Masters Projects by an authorized administrator of W&M ScholarWorks. For more information, please contact scholarworks@wm.edu.
DEVELOPMENT OF THE CHAE OPTIMAL SUPERVISION ENVIRONMENT TEST

A Dissertation
Presented to
The Faculty of the School of Education
The College of William and Mary in Virginia

In Partial Fulfillment
Of the Requirements for the Degree
Doctor of Education

by
Ki Byung Chae
July 2013
Dedications

I praise God, my Lord Jesus Christ, for all the work he has completed in me. For your steadfast love is great above the heavens; your faithfulness reaches to the clouds.

Special thanks to my wife and best friend, Ja Young, for her many prayers, unconditional love, and constant support. You are my love, my beautiful one.

To my two kids, Yuna and Yuri, for their beautiful smiles.

To my father and mother for showing confidence in me, supporting me to fulfill my dream, and being the best parents in the world.

To my sister for providing encouragement and support.

To my father and mother in laws for their constant prayers and care for my family.

To Manna Community Church for many prayers and joyful memories.

To my fellow classmates at the College of William and Mary for providing continuous support and encouragement. Special thanks to my doctoral cohort, Deneen, Derek, Joey, and Morgan.

And finally, to my esteemed professors at the College of William and Mary for providing the support and establishing a high standard for counseling, teaching, and research.
### Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>CHAPTER ONE</strong></td>
<td>1</td>
</tr>
<tr>
<td>Importance of Clinical Supervision</td>
<td>1</td>
</tr>
<tr>
<td>Supervision as a Developmental Process</td>
<td>3</td>
</tr>
<tr>
<td>The Supervision Environment</td>
<td>10</td>
</tr>
<tr>
<td>The Optimal Supervision Environment</td>
<td>11</td>
</tr>
<tr>
<td>The Emotional Environment</td>
<td>11</td>
</tr>
<tr>
<td>The Learning Environment</td>
<td>15</td>
</tr>
<tr>
<td>The Power Environment</td>
<td>19</td>
</tr>
<tr>
<td>A Lack of Research in Supervision</td>
<td>22</td>
</tr>
<tr>
<td>Supervision Models Related to Supervision Environment</td>
<td>24</td>
</tr>
<tr>
<td>Integrated Developmental Model (IDM)</td>
<td>24</td>
</tr>
<tr>
<td>Supervisory Working Alliance Model</td>
<td>27</td>
</tr>
<tr>
<td>Instruments Related to Supervision Environment</td>
<td>30</td>
</tr>
<tr>
<td>Revised Barrett-Lennard Relationship Inventory (RI)</td>
<td>30</td>
</tr>
<tr>
<td>Working Alliance Inventory-Trainee (WAI-T)</td>
<td>32</td>
</tr>
<tr>
<td>Supervisory Working Alliance Inventory (SWAI)</td>
<td>34</td>
</tr>
<tr>
<td>Role Conflict and Role Ambiguity Inventory (RCRAI)</td>
<td>36</td>
</tr>
<tr>
<td>Supervisory Relationship Questionnaire (SRQ)</td>
<td>37</td>
</tr>
<tr>
<td>Need for a New Measure</td>
<td>39</td>
</tr>
<tr>
<td>Theoretical Considerations</td>
<td>40</td>
</tr>
<tr>
<td>Constructive Developmental Theory</td>
<td>40</td>
</tr>
<tr>
<td>Attachment Theory</td>
<td>48</td>
</tr>
<tr>
<td>Purpose of the Study</td>
<td>59</td>
</tr>
</tbody>
</table>
Reliability 94
Validity 95
Implications to Counselor Education and Supervision 98
Limitations 100
Directions for Future Research 104
Conclusion 101
Bibliography 107
Appendix 131
Vita 142
List of Tables

1. Table 2.1 COSET Gender by Race Sample Frequencies  
   
2. Table 2.2 Distribution of COSET Sample Professional Roles  
   
3. Table 2.3 Distribution of COSET Sample Supervision Experience  
   
4. Table 3.1 Pattern Matrix and Communalities for the Three-Factor Model of  
   COSET  
   
5. Table 3.2 Means, Standard Deviation, and Item-Scale Correlations of  
   COSET  
   
6. Table 3.3 Means and Intercorrelations for EE, LE, and PE
# List of Figures

1. Figure 2.1 COSET Blueprint   63
2. Figure 3.1 Three-Factor Model of COSET   85
3. Figure 3.2 One-Factor Model of COSET   86
DEVELOPMENT OF THE CHAE OPTIMAL SUPERVISION ENVIRONMENT TEST

Abstract

The review of current supervision models and instruments revealed a crucial need for a valid, reliable instrument that assesses the quality of the supervision environment as a venue for promoting counselor development. Therefore, the primary purpose of this study was the construction and initial validation of the Chae Optimal Supervision Environment Test (COSET). The five phases of scale development provided preliminary evidence of reliability and validity for the COSET. The COSET was administered to 93 counselor educators and clinical supervisors. Results indicated that the 15 item COSET possesses three factors: Emotional Environment, Learning Environment, and Power Environment. Reliability data also revealed that the COSET and its factors have adequate evidence of internal consistency. A three-factor COSET model demonstrated a good model fit using a confirmatory factor analysis. The results are largely supportive of the COSET as a scale to assess supervisors’ creation of optimal supervision environments. Implications for the study and training of counselor supervisors were highlighted and suggestions for future research were reviewed.

Ki Byung Chae

School of Education: Counselor Education and Supervision

THE COLLEGE OF WILLIAM AND MARY IN VIRGINIA
CHAPTER 1
INTRODUCTION

Importance of Clinical Supervision

Supervision has been found to be a critical element in the training and development of professional counselors (Bernard & Goodyear, 2009). The supervisory relationship often is the most formative relationship that novice counselors experience as they develop professional identities (Riggs & Bretz, 2006). Furthermore, supervision consistently promotes counselors' growth and development so that they satisfy the standards of the profession and ensure therapeutic effectiveness (Barrett & Barber, 2005; Bernard & Goodyear, 2009; Holloway & Neufeldt, 1995).

Bernard and Goodyear (2009) noted that: “Practice alone is an insufficient means by which to attain competence: Unless it is accompanied by the systematic feedback and guided reflection that supervision provides, supervisees may gain no more than the illusion that they are developing professional expertise” (p. 4). With this in mind, they defined supervision as follows:

Supervision is an intervention provided by a more senior member of a profession to a more junior member or members of that same profession. This relationship is evaluative and hierarchical, extends over time, and has the simultaneous purposes of enhancing the professional functioning of the more junior person(s); monitoring the quality of professional services offered to the clients that she, he, or they see; and serving as a gatekeeper for those who are enter the particular profession (Bernard & Goodyear, 2009, p. 8).
Despite the potential benefits of supervision, it appears the experience of supervision can also be negative and even damaging for supervisees (e.g., Ellis, 2001; Gray, Ladany, Walker, & Ancis, 2001; Greer, 2002; Jernigan, Green, Helms, Perez-Gualdrón, & Henze, 2010; Ladany, Lehrman-Waterman, Molinaro, & Wolgast, 1999; Magnuson, Wilcoxon, & Norem, 2000). For example, Gray et al. interviewed 13 psychotherapy trainees to explore their experiences in “counterproductive” supervision events. The researchers defined a counterproductive event as “any experience that was hindering, unhelpful, or harmful in relation to the trainee’s growth as a therapist” (p. 371). Participants all reported counterproductive experiences, including: supervisors dismissing trainees’ thoughts and feelings, lacking empathy, and inappropriately self-disclosing. Most perceived counterproductive events were attributed to supervisors not attending to their trainees’ thoughts and feelings. Trainees reported changing their behaviors toward their supervisors after those experiences, most commonly by repressing disclosure. Ellis (2001) also reported negative supervision experiences, such as: the lack of supervisor empathy and support, supervisees feeling unsafe and withdrawing from the supervisory relationship, supervisees developing self-doubts and blaming themselves, and supervisees experiencing diminished self-efficacy as professionals. Nelson and Friedlander (2001), who interviewed 13 master- and doctoral-level trainees, reported that “bad supervisors” were viewed by trainees as being “remote and uncommitted to establishing a strong training relationship” (p. 387). As a result of perceived bad supervision, some of the trainees reported experiencing long-lasting self-doubt
and extreme stress. To prevent such perceived bad and even harmful supervision experiences, it is essential for supervisors to identify and understand factors that contribute to a safe, positive supervision environment.

Supervision as a Developmental Process

Cognitive developmental theorists (e.g., Dewey, 1963; Hunt, 1971; Kohlberg, 1981; Piaget, 1970) posited that the development of cognitive, internal structures for constructing meaning from experiences emerge as the result of interactions between individuals and their environment and the restructuring of psychological schema to meet changing environmental demands. Cognitive structures determine how individuals make sense of their environment and how they react to it (Sprinthall, Peace, & Kennington, 2001). Cognitive developmental theories are based on a number of foundational assumptions, including: (a) individuals process experiences through cognitive structures; (b) cognitive structures are organized in an invariant, hierarchical succession of stages from the less complex to the more complex; (c) stages consist of distinct, qualitative differences in how individuals make meaning from their experience; (d) development is not automatic; and (e) behaviors can be determined and predicted by a particular stage of development (Reiman & Thies-Sprinthall, 1998). Blocher (1983) proposed a central hypothesis of cognitive development as follows: “An individual’s perceptions of others tend to develop in the direction of greater complexity, decreasing stereotypy, and greater ability to integrate discordant or inconsistent information about the behavior of others” (p. 27). According to cognitive developmental theories, new experiential challenges requiring higher
levels of “meaning-making” create psychological dissonance, which, in turn, causes individuals to revise their current cognitive structures, resulting in psychological growth.

The application of a developmental approach to supervision and counselor development is not a new concept. Hogan (1964) was among the first theorists to describe supervision as a process focusing on a counselor’s developmental level. He argued that supervisors need to match their supervision interventions to the needs of their supervisees. Since Hogan’s insights, a lot of research has focused on the application of developmental theories in counselor development and supervision (e.g., Bernard & Goodyear, 2009; Fong, Borders, Ethington, & Pitts, 1997; Foster & McAdams, 1998; Granello, 2002, 2010; Lovell, 2002; McAuliffe & Lovell, 2005; Watkins, 1997). Nelson, Barnes, Evans, and Triggiano (2008), for example, reported on the experience of a supervisor in which supervisee resistance was framed as a developmental process:

I think that there is a developmental stage when a supervisee wants to disagree with their supervisor and needs to, when they’re really sort of testing their own frame of reference… and that we need to support that…and we don’t want them to be sponges, we really want them to go their own direction (p. 178).

Recent studies have applied developmental theories to counselor education and conclude that counseling students typically progress through a series of developmental stages as they advance through their counseling programs (Fong et al., 1997; Granello, 2002). Granello and Hazler (1998) argued that counselor
educators need to present material and skill development based on the level of the counseling students. Pearson (2000) suggested that the supervisee’s developmental level can influence preferences toward such factors as: balancing support with challenging interventions, desiring structure, focusing on teaching and skill development, and exploring countertransference issues. Furthermore, Handley (1982) recommended that supervisors and supervisees to be aware of their cognitive styles early in the supervision process that allows supervisors to anticipate potential problem areas that reflect the interaction of cognitive styles, leading to more productive and positive experiences in supervision.

Swanson and O’Saben (1993) surveyed 57 counselor and psychologist trainees to examine the relationships between trainees’ cognitive styles, program membership, amount of practicum experience, and perceived needs and expectations for supervision. The Myers-Briggs Type Indicator (MBTI; Myers & McCaulley, 1985) was used to measure participants’ cognitive style, the trainee version of the Supervisor Perception Form (SPF-T; Heppner & Roehlke, 1984) was used to assess participants’ expectations of their supervisors, and the Supervisory Needs subtest of the Counselor Development Questionnaire (CDQ-SN; Reising & Daniels, 1983) was used to assess the perceived needs of trainees at each level of development during the supervision experience. Results from the Swanson and O’Saben study indicated that trainees’ perceived expectations and needs regarding the supervisory experience differed by program membership, amount of practicum experience, and cognitive styles. Specifically, they found that supervisees with less practicum experience demonstrated more openness to
learning, a greater need for direct observation of their therapy sessions, a stronger desire for their supervisor to influence their behaviors in therapy and provide concrete counseling techniques, and a greater desire for immediate access to their supervisor during crisis situations, as well as the availability of an intervention strategy from their supervisor. However, the authors noted that the SPF-T and CDQ-SN lacked empirical support for validity and reliability, and so questioned the meaningfulness of the conclusions drawn. Also, the study did not directly assess variables related to actual supervision, such as the supervisory relationship or supervision outcome, thus suggesting the need for additional research. Despite the limitations, the results of the Swanson and O'Saben study support the notion that supervisees, at varying levels of counseling experience and differing cognitive styles, have different expressed expectations and needs for supervision. By understanding of their supervisees’ counseling experiences and cognitive styles, supervisors can more likely anticipate relevant issues that comprise the supervision process.

Granello (2010) examined the extent to which the number of years of counseling experience in the field can predict levels of cognitive complexity in a sample of licensed counselors. Participants in this study were 122 licensed counselors from one Midwestern state. The researcher used the Learning Environment Preferences (LEP; Moore, 1989) to assess participants’ levels of cognitive development. The LEP consists of 65 items and includes five domains related to epistemology and approaches to learning. The researcher modified the original scale by changing the words of the items to reflect experiences related to
continuing education and learning within the field of counseling rather than general classroom learning. The five modified-LEP domains are: (a) beliefs of the nature of knowledge of counseling, (b) role of the instructor or workshop presenter for continuing education purposes, (c) role of the participant in his or her own continuing education, (d) atmosphere of the learning environment, and (e) role of evaluation. Results of a stepwise regression analysis indicated that among all the predictor variables (i.e., years in the counseling profession, years as a practicing counselor, highest degree, age, gender, and race), years in the counseling profession emerged as the most significant contributor to counselor cognitive complexity \( (R = .34, R^2 = .11, p < .001) \). The researcher also reported that among counselors with the most experience, there were more counselors in the higher stages of Perry’s (1970) cognitive developmental model and fewer in the lower stages. The findings from this study provide direct evidence of an increasing complexity along the developmental path of professional counselors. However, the researcher noted that the study was cross-sectional, providing only preliminary conclusions, and recommended a larger scale longitudinal study that would provide a stronger database from which to examine the cognitive development of counselors in-training.

Collectively, research literature strongly supports the importance of an informed understanding of supervisees’ developmental differences and the effects of those differences on supervisory relationships and outcomes. A more in-depth understanding of supervisees’ differences can help counselor educators implement optimal supervision strategies for supervisees at different developmental levels.
(Fernando & Hulse-Killacky, 2005). Recently, Kegan's (1982) constructive developmental theory (CT) has been employed in understanding the complexity of counselor development (e.g., Eriksen, 2007, 2008; Grigoriu, 1998; McAuliffe 1993; Paul, 2008; Pratt, 1998). Applying Kegan's theory, Pratt (1998) investigated how individuals perceive critical issues in clinical work differently depending on their developmental level. Using a semi-structured interview format, the researcher interviewed 12 female psychologists with at least five years of clinical experience, ranging from five to 20 years. Six questions related to issues in clinical practice comprised the interview protocol. Issues underlying the interview questions were as follows: (a) responses to client manipulation; (b) dealing with counseling termination; (c) therapists' manner of dealing with dual relationships; (d) perceptions of therapeutic challenges; (e) perceptions about supervisory relationships; and (f) changes experienced as therapists. Data were analyzed according to the data analysis procedure of the Subject-Object Interview (Lahey, Souvaine, Kegan, Goodman, & Felix, 1988), which reflects Kegan's model and the qualitative analysis of emergent themes from questions. Results from the analyses revealed that supervisees demonstrated developmental differences in their meaning-making in how they responded to four of the questions: (a) manipulative clients, (b) termination, (c) changes, and (d) challenges. The most common reported by change supervisees was a perceived increased ability to set limits and maintain boundaries. Supervisees' reports generally supported Kegan's approach for explaining therapists' understanding of their clinical practice. According to Kegan's theory, these outcomes reflect
developmental movement from stage three (i.e., interpersonal) to stage four (i.e., institutional). In fact, all of the therapists’ perceptions of change were in line with developmental expectations, providing support for the hierarchical and invariant sequence to Kegan’s proposed constructive development. However, regarding the methodology of this study, the researcher questioned the efficacy of Kegan’s model due to the difficulty of coding and processing data for the Subject-Object Interview. The researcher suggested that a more standardized method for coding in Kegan’s model would be an improvement for future research.

Kegan’s theory focuses on understanding the “deep structures” (Rogers & Kegan, 1991, p. 105) of knowing that underlie the development of our self-conception and capacity for relating to others. The “structures” are viewed as a framework by which individuals understand the world, including their self-awareness and perceptions of others. Kegan (1982) described development as a process in which individuals construct and reconstruct personal meaning over the life span. Like other cognitive developmental theorists, Kegan posited that psychological growth takes place as a result of the interaction between the person and the environment through the processes of assimilation and accommodation. As new experiences challenge the limits of an individual’s ability to make meaning using his or her current meaning-making system (i.e., assimilation), a new system for understanding the self and world is constructed (i.e., accommodation). As individuals progress through developmental stages, they achieve increasingly more expansive, open, and inclusive understanding of themselves and the world. Therefore, by understanding the supervisees’ meaning-
making system, supervisors might better be able to construct an effective supervisory relationship.

**The Supervision Environment**

In an effort to enhance supervision, considerable research interest has focused on the importance of matching supervisees’ developmental levels with appropriate supervisory conditions, typically referred to as the “supervision environment” (Bernard & Goodyear, 2009). In general, individuals progress through a series of qualitatively distinct levels of complexity with which experience is organized and understood (Kegan, 1982). According to Kegan (1982), development from one level to another occurs through an interaction between individuals and their environments, which influences many dimensions of individuals’ experiences, including cognitive, affective, interpersonal, and intrapersonal experiences.

When supervisors match supervisory interventions to their supervisees’ current developmental level and then mismatch their interventions to their supervisees by relating from the next developmental level, this approach optimize the supervisory environment. Stoltenberg (1981) noted that the optimal supervision environment is one in which there is a mismatch in challenge of about one-half step beyond the supervisee’s current level of functioning. This optimal mismatch extends the supervisees’ thinking but does not overwhelm the supervisees’ thinking with more information that they can handle. Borders (1998), applying the framework of ego development, suggested that for supervisees to transition to a higher level of ego development within the context of supervision,
the supervisor must be functioning at least one ego developmental stage higher than their supervisees.

**The Optimal Supervision Environment**

Supervisors, therefore, clearly have a responsibility to create an optimal supervision environment, making adjustments as needed, based on the developmental needs and characteristics of supervisees. Drawing from the current supervision literature (e.g., Barrett & Barber, 2005; Borders, 1998; Dickson, Moberly, Marshall, & Reilly, 2011; Ellis, 2001; Falender et al., 2004; Fitch, Pistole, & Gunn, 2010; Foster, Lichtenberg, & Peyton, 2007; Holloway, & Neufeldt, 1995; Kilminster & Jolly, 2000; Ladany et al., 1999; Ladany, Lehrman-Waterman, Molinaro, & Wolgast, 1999; Magnuson et al., 2000; Neswald-McCalip, 2001; Palomo, Beinart, & Cooper, 2010; Riggs & Bretz, 2006; Stoltenberg, 2005; Wheeler & Richards, 2007; White & Queener, 2003; Worthen & McNeill, 1996), it appears that an optimal supervision environment is composed of three core elements: (a) the emotional environment, (b) the learning environment, and (c) the power environment.

**The Emotional Environment (EE).** In supervision environments, the supervisory relationship serves as a bridge from where supervisees have been to where they are going (Eriksen, 2008). Eriksen (2008) suggested that supervisees should be able to stand at any point on the bridge and feel well supported. The quality of the supervisory relationship is a critical element that produces positive outcomes in supervision and promotes counselor development (Bernard & Goodyear, 2009; Worthen & McNeill, 1996). The fact that a positive and
productive supervisory relationship is critical to positive counseling outcomes is well documented (Black, 1988; Nelson, Gray, Friedlander, Ladany, Walker, Melincoff, 2001; Palomo et al., 2010; Ramos-Sanchez et al., 2002; Webb & Wheeler, 1998; Worthen & McNeill, 1996). Shanfield, Mohl, Matthews, and Hetherly (1992), for example, examined 34 psychotherapy supervisors using the Psychotherapy Supervisory Inventory (PSI; Shanfield, Mohl, Matthews, & Hetherly, 1989) to explore patterns of supervisory behavior. Participants were asked to record their supervision sessions, and the researchers used the PSI to rate supervisors’ behaviors. Thirteen scales of the inventory served as predictors of supervisees’ perceived excellence of supervisors, as a measure of supervisee quality. Using a stepwise multiple regression, the researchers found that empathy towards the supervisee was the most powerful predictor of all variables investigated for judging effective supervision, and that focusing on the supervisee was also a significant predictor. Although the study noted some limitations (e.g., lack of data on supervision outcomes, lack of data on supervisors’ or supervisees’ perspective, and raters’ biases), the study outcomes suggest that empathy toward and focus on supervisees are important elements for supervision effectiveness and for enhancing supervisees, trust in supervisory relationships.

Studies have demonstrated a strong association between supervisors and supervisees’ emotional bond and various supervision outcomes (Ellis, 2010; Ladany, 2004; Ladany, Ellis, & Friedlander, 1999; White & Queener, 2003). Ladany et al. (1999) investigated the relationships between supervisory alliance, supervisee self-efficacy, and supervisees’ satisfaction with supervision. The
researchers studied 107 counselor and psychologist trainees, using the Working Alliance Inventory-Trainee version (WAI-T; Bahrick, 1990) to assess supervisory working alliance; the Self-Efficacy Inventory (SEI; Friedlander & Snyder, 1983) to assess trainee self-efficacy; and the Trainee Personal Reaction Scale-Revision (TPRS-R; Holloway & Wampold, 1984) to assess trainees’ perceived satisfaction with supervision. Although the supervisory alliance did not predict changes in supervisees’ self-efficacy, a strong emotional bond was predictive of supervisees’ satisfaction with supervision. That is, as the emotional bond between supervisor and supervisee increased in strength, supervisees perceived their supervisors’ personal qualities and performance more positively, they perceived their own behaviors in supervision more positively, and they perceived a higher level of comfort in supervision. Because of the limitation stemmed from the threats to validity inherent in ex post facto designs (i.e., a non-experimental research design in which preexisting groups are compared on some dependent variable), the study was unclear whether positive changes in the emotional bond led to greater satisfaction with supervision or whether greater satisfaction with supervision led to positive changes in emotional bond. Furthermore, the researcher suggested that the supervisory process should be examined from other perspectives (e.g., observers or supervisors).

According to Watkins (2010), the supervisor establishes the relationship as a container or holding environment (Winnicott, 1975) to create a safe space for the supervisee, wherein trust, consistency, and dependability permeate every facet of the supervisory relationship. Specifically for novice supervisees, supervisors
are suggested to hold the supervisee more tightly (opposed to later) and to serve a much needed compensatory function (Watkins, 2010). Watkins suggested that when the supervisee feels anxiety, the supervisor should provide soothing; when supervisees have doubts, the supervisor should provide reassurance, and; when the supervisee lacks direction, the supervisor should provide guidance.

White and Queener (2003) examined the relationship between supervisors’ and supervisees’ self-reported abilities to make healthy adult attachments in relationships, social provisions (i.e., social network), and their perceptions of supervisory working alliance. The researchers examined 67 supervisors working in professional and academic settings and 67 supervisees recruited from three Midwestern university programs. Supervisory working alliance was measured using the Supervisory Working Alliance Inventory (SWAI; Efstation, Patton, & Kardash, 1990). The quality of the supervisors’ and supervisees’ social network was assessed using the Social Provision Scale (SPS; Cutrona & Russell, 1987). Supervisors’ and supervisees’ perceived ability to develop healthy adult attachments and relationships with others was measured using the Adult Attachment Scale (AAS; Collins & Read, 1990). Results indicated that the supervisors’ ability to create secure adult attachments and social provision (i.e., social network) were predictive of both supervisees’ and supervisors’ perceptions of the supervisory working alliance. On the other hand, supervisees’ adult attachments and social provisions were not significant predictors of either supervisees’ or supervisors’ perceptions of the working alliance. The researchers concluded that the supervisor’s abilities to form close attachments and to feel
intimate in relationships are more predictive of the supervisory alliance than the
same characteristics brought to the supervisory relationship by supervisees.
However, the researchers noted some limitations of this study. Because the study
used mostly female, master's level counseling students at only three universities,
researchers cautioned the generalizability of the findings. Also, due to the ex post
facto design used in this study, causal inference cannot be made (e.g., supervisors’
ability to make attachments and social provisions caused working alliances to be
varied). Despite these limitations, this study offers further empirical support for
the link between the supervisors’ ability to create a supportive supervisory
relationship and the perceived effectiveness of supervision environment. In
summary, recent studies provide support for the notion that supervisors’ ability to
create an emotional bond and a secure attachment relationship is critical to create
an effective supervision environment.

The Learning Environment (LE). To provide effective supervision,
Borders (1989b) suggested that supervisors must consider their supervisees as
"learners" and of themselves as "educators" who create learning environments (p.
6). She stated that competent supervisors are not only competent counselors but
also skilled educators who impart their counseling knowledge and skills by
matching supervision interventions according to their supervisees' cognitive
developmental levels. In addition, research literature shows the importance of
creating a learning environment that supports and challenges the supervisees’
cognitive developmental level (e.g., Borders, 1989a; Borders & Fong, 1989;

Studies have supported the assumption that cognitive complexity increases during supervised counseling practice (e.g., Granello, 2002, 2010). The process of developing a comprehensive understanding of clients and case conceptualization is complicated and requires that the counselor have advanced cognitive processing abilities (Blocher, 1983; Welfare & Borders, 2010). Supervision can serve as an optimal environment for counselors to enhance their cognitive capacity.

Borders et al. (1986) studied the relationship of 63 counseling practicum and intern students' levels of ego development and levels of experience with perceptions of their clients. The researchers used the Sentence Completion Test of Ego Development (SCT; Loevinger & Wessler, 1970) to assess cognitive complexity, termed "ego development level," and the Repertory Grid Technique (Kelly, 1955; Neimeyer & Neimeyer, 1981) to measure client cognitions. The researcher specifically wanted to know what the impact of student ego developmental level and experience level has on: (a) the complexity of thoughts about clients and (b) content of thoughts about clients. Results indicated that students' experience levels did not differentiate between their client perceptions but found an effect on ego developmental level. The researchers suggested that because of the small sample size, the overall impact of student ego developmental level on complexity and client cognitions might be inconclusive. On the other hand, the researchers found some differences among thought content across
students’ ego levels. They described those students at lower ego levels using more simplistic, concrete descriptors, whereas those at higher ego levels using more sophisticated and interactive descriptors. The researchers stressed the importance of evaluating students’ thoughts for both complexity and content to find subtle effects. This study provides support for promoting counselor cognitive development in training and supervision.

Borders (1989a) conducted a study on 27 practicum students to investigate the influence of ego development on in-session cognitions of supervisees at the same level of counseling experience. The researchers used the SCT to measure the overall cognitive complexity of the students. The students had varied levels of cognitive complexity despite having the same level of counseling experience. Students recorded their actual counseling sessions and reviewed the tapes immediately following their counseling sessions. As the students watched the tape, they verbalized their thoughts and feelings. These recall sessions were taped, transcribed, and reviewed to code the retrospections. Results indicated that counseling students with higher levels of cognitive complexity (i.e., higher level of ego development) reported significantly fewer negative thoughts about clients and their performance and were better able to remain objective and neutral in the counseling sessions. Despite the small sample size and the limited range of ego developmental level, this study implies that some counseling tasks are performed better by individuals at higher levels of ego development, regardless of training and experience level.
A longitudinal study of counselor cognitive development was conducted by Fong et al. (1997). The researchers assessed 43 counseling students' cognitive development five times from the beginning to end of their counseling program: (a) at the start of the program; (b) at the completion of the first semester; (c) 3 semester-hour counseling-skill training course; (d) at the end of practicum; and (e) at the end of their final internship. The SCT was used to assess level of student cognitive development. The Stress Appraisal Scale (SAS; Carpenter & Suhr, 1998) was used to measure students' own thoughts and feelings about providing counseling services. The researchers reported that students' levels of ego development did not change over the course of the program. They assumed that the SCT is too general, and the levels are too broad to identify the changes in cognitive complexity that occurs during a counseling program. Despite these limitations, they reported that students with higher levels of cognitive development used more complex and effective verbal skills, had more confidence in their work, and found counseling less difficult. The authors suggested additional longitudinal studies of counselor cognitive development and cited the need for more specific measures of counselor cognitions.

In summary, these studies support the importance of providing a learning environment in train and supervision that promotes counselor cognitive development. Counselors at higher levels of cognitive development are better able to formulate a thorough, objective understanding of the client and communicate effectively and confidently in the counseling sessions. Research supports the notion that supervision is the ideal setting to promote counselors'
cognitive complexity. Thus, supervision can be an essential component of promoting cognitive complexity. Supervisors' ability to understand and promote supervisees' cognitive development by creating an effective learning environment appears to be a critical element of the optimal supervision environment.

The Power Environment (PE). One important aspect of supervision that is significantly different from counseling is that of power inequality. Supervisors in counselor training programs and the field of counseling are responsible for evaluating trainees' professional performance and monitoring the quality of the supervisory relationship (ACA, 2005; CACREP, 2009). While therapeutic relationship enhancement techniques (e.g., empathy, immediacy, self-disclosure, confrontation, and respect) translate well in supervision, evaluation adds layers of complexity to the supervisory relationship (Pearson, 2000). This evaluative component of supervision grants supervisors an important source of interpersonal influence (Bernard & Goodyear, 2009).

Even within the ideal supervisory relationship, evaluation and discussion of supervisees' personal challenges are inherent qualities of supervision that can provoke anxiety (Pearson, 2000). Most supervisees are required to be vulnerable and self-disclose their challenges in professional and personal growth to the same supervisors that evaluate them. Such expectations naturally generate tension for both supervisees and supervisors, leading to potential relational conflicts (Ladany, Friedlander, & Nelson, 2005; Nelson & Friedlander, 2001).

Because the supervisors are in a position of greater authority, the ability of supervisees to communicate their needs may be hindered. Jacobs (1991) noted
that "because students are emotionally vulnerable in the context of their supervision, they are in a poor position to advocate for themselves should the boundaries of that relationship break down" (p. 133). Ladany, Hill, Corbett, and Nutt (1996) examined 108 therapists in training to investigate the nature and content of supervisees' nondisclosure and the reasons for different types of nondisclosure. They found that 90% of the supervisees experienced a negative reaction to a supervisor, and most supervisees (97.2%) did not disclose their negative experiences in supervision due to the consciousness of supervisor's power or authority, impression management (i.e., fear of evaluation), and fear of retaliation. They concluded that "a good alliance with the supervisor is important if the supervisee is to feel comfortable revealing significant information, particularly negative reactions to the supervisor" (p. 21). However, the researchers cautioned that a causal link cannot be inferred from this correlational study. Also, the findings of this study reflect only the supervisees' perspectives and supervisors may offer alternative explanations for supervisees' nondisclosures. Despite these limitations, this study provides important indications that supervisees are more inclined to not disclose counterproductive events in supervision due to the power differential, placing supervisors in a difficult position to receive adequate feedback about their supervision performance.

In a qualitative study, Henderson, Cawyer, and Watkins (1999) interviewed five supervisors and ten supervisees participating in a practicum course to explore their perceptions of effective practicum supervision and to determine which factors in supervision the participants considered most important.
Results indicated that supervisees viewed the effectiveness of supervision to be a function of their supervisors' general level of knowledge and clinical experience, capacity to facilitate learning, and ability to offer constructive evaluation of clinical performance. Supervisees also identified relationship factors such as trust, approachability, respect, and attentiveness as important supervisor characteristics. While the supervisors reported similar relationship factors, they also emphasized the importance of attending to counselor development, ethics, and adaptability. Most importantly, the primary point at which the supervisees' and supervisors' responses differed from one another was in the area of evaluation. While supervisees noted that supervisors' ability to provide constructive evaluation of supervisees' performance as an important element of effective supervision, evaluation did not emerge as an important feature of the supervisors' perceptive of effective supervision. The researchers added that supervisors' abilities to communicate clearly, to provide a balance between support and constructive criticism, and to consider the supervisee's vulnerability when providing feedback were highlighted by the supervisees as important factors of constructive evaluation.

In another qualitative study, Nelson, Barnes, Evans, and Triggiano (2008) interviewed 12 supervisors recognized by professional peers as highly competent to examine their experiences of conflict in supervision and their reliable strategies for conflict management. The researchers highlighted important characteristics of competent supervisors such as openness to conflict, willingness to acknowledge their weakness, a focus on establishing a strong supervisory alliance, discussing
evaluation early on, and gaining regular feedback from their supervisees. Strategies to effectively manage conflicts included “contextualizing conflicts in light of developmental and environmental factors, seeking consultation with colleagues, self-coaching, processing conflicts, accentuating supervisee strengths, interpreting parallel processes, and withdrawing from supervisee dynamics” (Nelson et al., 2008, p. 172). This study further supports the need for supervisors to understand the hierarchical, evaluative nature of the supervisory relationship and to create an evaluative environment that promotes counselor development.

A Lack of Research in Supervision

Despite the essential role of the supervision environment in the development of counselors, there is a lack of research evaluating the quality of clinical supervision and a critical need for more structured and methodologically sound research (Bernard & Goodyear, 2009; Kilminster & Jolly, 2000; Wheeler & Richards, 2007). For example, Goodyear, Bunch, and Claiborn (2005) searched for supervision articles that had been published during the previous five years in psychology journals and found only 22 empirical studies. Ellis, Ladany, Krengel, and Schult (1996) examined 144 empirical studies in clinical supervision and found several studies with unchecked Type I and Type II error rates, moderate effect sizes, and lack of attention to hypothesis validity. Furthermore, the quality of existing supervision research is reported as “substandard” (Ellis & Ladany, 1997, p. 492), suggesting that few conclusions can be legitimately drawn from it to inform the preparation of supervisors.
Although, many supervisors and educators have applied supervision models to explain the supervisory relationship or the supervision environment, there is lack of empirical research that supports the use of those models. For example, one of the most well established counselor developmental models, the Integrated Developmental Model (IDM), has been criticized for lacking empirical testing (Ellis & Ladany, 1997). Stoltenberg (2005), the developer of this model, even stated that: “The field would benefit from more specific attention being paid to testing existing supervisory theory, including the IDM” (p. 862).

One of the main reasons for the lack of supervision research is the concurrent lack of reliable supervision instruments. Watkins (1998) stated that: “Research is only as good as the measurement tools and procedures that are used for assessment and evaluation” (p. 94). He argued the need of more valid, reliable, and supervision-specific measurements to advance research efforts. Many measurements, such as the Supervisory Levels Questionnaire-Revised (SLQ-R; McNeill, Stoltenberg, & Romans, 1992) or the Supervisory Working Alliance Inventory (SWAI; Bahrick, 1990) have been criticized for not having adequate psychometric soundness. Goodyear and Bernard (1998) have pointed out that previous measurement development has placed too much emphasis on supervisee satisfaction, a variable that is not necessarily predictive of skill development or clinical service provision. In addition, Watkins (1998) criticized that many supervision measurements have been borrowed from psychotherapy research and have not necessarily been developed with the supervisory environment in mind. He also noted that many supervision measurements were created for the particular
study at hand and never used again. Although, his critique on supervision research and measurement was written more than a decade ago, a review of current supervision models and measurements suggests that a critical need still remains for the development and establishment of reliable, valid, criterion measurements to guide supervision research.

**Supervision Models Related to Supervision Environment**

**Integrated Developmental Model (IDM)**

The Integrated Developmental Model (IDM; Stoltenberg, McNeil, & Delworth, 1998) is the best known and most widely used stage developmental model that focuses on the supervision environment (Bernard & Goodyear, 2009). The IDM is useful in conceptualizing the process by which counselors in training and practice increase their competency within various domains of professional practice (Stoltenberg, 2005). The IDM describes counselor development as moving through four stages, each of which is characterized by changes on three overriding structures—self-other awareness, motivation, and autonomy. These stages provide markers in assessing professional growth (Stoltenberg et al., 1998).

In the first stage, the supervisee has limited training, and is characterized by high motivation to learn, strong dependence on the supervisor, and high self-focus. The second stage includes fluctuating confidence resulting in varied levels of motivation toward autonomy and dependence but greater ability to focus on clients. The third stage involves an emerging personalized approach to practice, consistent motivation, movement towards autonomy, and the ability to focus on the client while remaining self-aware. The fourth stage characterized the ability
to function at stage three across a variety of domains while using a flexible, personalized approach.

According to the IDM, to accommodate the different developmental needs of supervisees, supervisors need to change their supervision interventions. For example, the IDM emphasizes that during the initial stages of supervision the supervisee should be offered significant structure, direction, and support to promote development. As a supervisee gains some experience, expertise, and confidence, the supervision environment moves toward less structured, non-directive supervision, and more challenges are assigned. The supervision environment should generally reflect a decrease in the amount of structure provided by the supervisor as supervisees develop. High levels of early structure in supervision should have the effect of helping control supervisees’ anxiety and give them direction for exploring and understanding the intervention process. As their skill levels and understanding increase, they are better able to take more responsibility for their learning and growth in the supervision context. This requires less external structure in the supervision environment.

Although the central assumptions and tenets of the IDM have received some support in literature (Worthington, 1987), other studies have challenged them. For example, Barrett and Barber (2005) argued that the IDM model fails to directly draw a link between the influence of personal development or maturity (cognitive and emotional development) and professional development. The researchers asserted that an individual’s cognitive and emotional development is a different developmental track—separate from a trainee’s professional
development yet influencing it. More specifically, Stoltenberg et al. (1998) suggested that an individual's level of personal maturity may serve as a ceiling, preventing some trainees from negotiating more advanced professional development in a timely or efficient manner. However, the IDM allegedly fails to account for this influence, thus neglecting the complexity of interpersonal supervisory relationship (Barrett & Barber, 2005). Without considering the trainee's level of personal maturation, supervisors may erroneously expect a trainee to learn and utilize skills beyond their current developmental level.

Holloway (1987) challenged the validity of the IDM by arguing that the model is flawed in its lack of recognition of other important variables that may account for trainees' behaviors and perceptions. Specifically, she noted the lack of attention given to the characteristics that trainees bring to the training process and the fact that the trainees' developmental stage could be influenced by the content or the quality of a particular problem. Similarly, Tracey, Ellickson, and Sherry (1989) examined the preference of beginning and advanced trainees for structured supervision as moderated by aspects of their personality and the content of supervision (crisis vs. non-crisis material). They found that advanced trainees with high reactance (i.e., high need to resist structure) preferred less structured supervision than other advanced trainees with low reactance. In non-crisis situations, beginning trainees preferred structured supervision, whereas more advanced trainees preferred less structure; however, all trainees preferred structured supervision in crisis situations regardless of their level of experience or reactance.
Supervisory Working Alliance Model

The working alliance model (Bordin, 1983) is used as a conceptual framework to explain the supervisory relationship. It refers to the collaboration of the supervisee and supervisor that facilitates change in the supervisee through a mutual agreement on the goals and tasks of supervision and through a strong emotional bond. Bordin (1983) associated the supervisory relationship with the therapeutic alliance found in counseling, in which the therapist and client work together to form and attain goals. The supervisory working alliance model is composed of three elements: (a) mutual agreement on the goals of supervision; (b) mutual agreement on the tasks necessary to attain the set forth goals of supervision; and (c) an emotional bond between supervisor and supervisee, which is described through interactions such as caring, trust, and attraction. This model suggests that relational bonds develop as a result of working together on a common task to achieve shared goals or on the basis of shared emotional experiences.

Several studies indicate that the supervisory working alliance is associated with significant supervisory outcomes. For example, Ladany et al. (1999), as mentioned earlier, examined 107 counselor trainees and investigated the extent to which changes in trainees' perceptions of the three components of the supervisory working alliance (i.e., goal, task, and bond) are related to changes in two supervisory outcomes (i.e., supervisee self-efficacy and satisfaction with supervision). They found that the supervisory alliance was not predictive of changes in the supervisees' self-efficacy, but the emotional bond was predictive
of supervisee satisfaction with supervision. This study suggests that as the emotional bond between the supervisor and supervisee became stronger over time, the supervisees perceived their supervisors’ personal qualities and performance more positively, their own behaviors in supervision more positively, and they experienced a higher level of comfort in supervision. The reverse was also found in that supervisees viewed themselves and their supervisors more negatively if the emotional bond did not increase over time. These findings were supported by White and Queener (2003) suggesting that supervisors' ability to create close, supportive relationships is predictive of both supervisees' and supervisors' perceptions of the supervisory working alliance.

Although research has consistently found that the supervisory working alliance is one of the most important elements in the process of supervision, there still seem to be some shortcomings with the working alliance model. First, there is no common measure of the supervisory working alliance (Efstation et al., 1990; Ladany et al., 1999). Some researchers have used either the Working Alliance Inventory-Trainee (WAI-T; Bahrick, 1990), and others have used the Supervisory Working Alliance Inventory (SWAI; Efstation et al., 1990) to measure the supervisory working alliance. However, it is important to note that the development of these two measurements was guided by different conceptual models. The 36-item WAI-T was a direct translation of Horvath and Greenberg’s (1986) Working Alliance Inventory which was based on Bordin’s (1979) model of the therapeutic working alliance. Although, the SWAI was also based on Bordin’s (1983) theory, it was also based on other theories such as Greenson’s
(1967), Robinson (1950), Gelso and Carter (1985), Patton (1984), and Pepinsky and Patton (1971). For example, Bordin delineated the concept of working alliance by identifying the three components of the alliance as goals, tasks, and bonds. On the other hand, Greenson viewed the working alliance from the psychoanalytic perspective which stresses the contribution of transference and countertransference issues to a therapeutic relationship. Also, Gelso and Carter extended Bordin’s definition suggesting that the working alliance is an emotional alignment that is fostered by the emotional bonds, agreement on goals, and agreement on tasks. Gelso and Carter rejected Bordin’s definition that the working alliance, itself, consists of goals, bond, and tasks. They argued that agreement on goals and tasks and the existence of a bond do not constitute the definition of the alliance, but influence and are influenced by the alliance. This theoretical difference between the two measurements may have result in constructing rather different concepts and identifying different factors. Second, there is little support for Bordin’s proposed three-factor model. Studies have suggested that high intercorrelations between the factors may indicate a single relational or bond factor or a two-factor bond and agreement model (Andrusyna, Tang, DeRubeis, & Luborsky, 2001; Ladany, Ellis, & Friedlander, 1999). Third, little is known about the personal variables that predict the ability to form quality relationship. The model pays little attention to the impact of the personal variables of supervisor and supervisee.
Instruments Related to Supervision Environment

Revised Barrett-Lennard Relationship Inventory (RI)

The Revised Relationship Inventory (RI) is a revised version of the Barrett-Lennard Relationship Inventory (Barrett-Lennard, 1962, 1969) that measures clients’ experience of five therapeutic facilitative conditions (i.e., regard, unconditionality, empathic understanding, congruence, and willingness to be known). Other than the last scale, added by Barrett-Lennard, these facilitative conditions are based on Rogers’ (1957) description of the conditions necessary and sufficient for constructive personality change. Rogers argued that the facilitative conditions are necessary in many different relationships, which include the supervisory relationship. The RI requires supervisees to indicate the degree to which they believe their supervisors provide particular facilitative conditions as reflected in statements on a 6-point Likert scale.

Originally composed of 92 items, the RI has been reduced to 85 items, and then further reduced to 64 items, omitting the willingness to be known scale because of its high correlation with congruence scale (Barrett-Lennard, 1969). Wiebe and Pearce (1973) developed a shorter form, in which they selected 32 items of the original 92, comprising four scales (excluding willingness to be known scale). A study by Dalton (1983) further refined this instrument, improving the reliabilities for the four scales ranging from .83 to .95. Schacht, Howe, and Berman (1988) then added the willingness to be known scale because of the potential importance of this factor in facilitating a closer identification with the supervisor. This resulted in a 40-item short form of the RI, with ten items in

30
each of the *regard*, *empathic understanding*, and *congruence* scales, and five each in the *unconditionality* and *willingness to be known* scales (Schacht et al., 1988). Psychometric data was obtained by Schacht et al. (1989) through a study in which they examined 152 participants who recalled their recent supervision experiences and rated the supervisors who they believed contributed the most and least to their therapeutic effectiveness. Total reliability for the 40-item RI was .92, while reliabilities for the scales ranged from .72 (*willingness to be known*) to .90 (*regard*). Schacht et al.'s results supported the notion that conditions comparable to unconditional warmth, empathy, and genuineness are necessary in order for supervision interventions to be effective.

Ellis and Ladany (1997) recommended the RI for research because of the sound psychometric evidence supporting it. However, the RI is based on a theory of personality change within therapy which may not adequately reflect the complexity of supervision environments. In addition, it offers limited directions for research and practice in supervision (Bernard & Goodyear, 2009). While, Schacht et al. (1998) suggested that the RI can be reworded to pertain to supervisees' past experience with their supervisors, a supervision measurement developed from a direct translation of terms used in theories of psychotherapy may not reflect the nature of supervision environment (Watkins, 1998). For example, one of the most important differences between therapy and supervision is the evaluative nature of supervision; however, the RI does not address this difference. Also, this instrument does not measure the developmental aspects of supervision environment such as supervisors' adjustment of their supervisory
style to match the supervisees’ level of development. Barrett and Barber (2005) noted that most negative supervisory experiences can be explained by the differing needs or experiences among trainees that are not addressed by the supervisor. Thus, it seems that the RI has several limitations to be used as a measurement that assesses the quality of supervision environment.

**Working Alliance Inventory-Trainee (WAI-T)**

The Working Alliance Inventory-Trainee (WAI-T; Bahrick, 1990) is a 36-item self-report instrument for measuring the quality of supervisory alliance based on Bordin’s (1983) pantheoretical conceptualization of the working alliance. According to Bordin’s (1983) model of the Therapeutic Working Alliance, the working alliance is described as the foundation needed in order for the client to consistently be receptive of the treatment. The three constituent components (e.g., tasks, bonds, and goals) in combination define the quality and strength of the alliance (Bordin, 1983). First, goals (outcomes) refer to the target of the intervention that that counselor and the client have mutual agreed upon. The clarity and mutuality of that agreement about goals contributes to the strength of the working alliance. Next, tasks refer to the counseling behaviors and cognitions that form the substance of the counseling process. The strength of the working alliance is also based on the mutual understanding by the supervisor and supervisee about the tasks that their shared goals impose on each other. Lastly, bonds comprise the complex network of positive personal attachments between the client and the counselor that include issues such as mutual trust, acceptance, and confidence. The WAI-T is a direct translation of Horvath and Greenberg’s
(1986) Working Alliance Inventory into supervision terms. For example, terms such as therapist and client have been changed to supervisor and trainee respectively. The WAI-T's three subscales—goals, tasks, and bonds—each contain 12 items and are measured by respondents rating statements about supervision on a 7-point Likert scale.

The WAI-T has been used in a range of supervision research studies (e.g., Ladany, Brittan-Powell, & Pannu, 1997; Ladany et al., 1999; Ladany & Friedlander, 1995; Ladany & Lehrman-Waterman, 1999; Walker, Ladany, & Pete-Carolan, 2007). However, similar to the RI, a primary problem with this instrument is the assumption that a measure of the therapeutic alliance will simply transfer to the supervisory setting without testing the assumption. As such, the instrument does not consider the evaluative and gate-keeping element of supervision. Also, the WAI-T and its underlying conceptual model seem to inadequately consider the complexity of the supervision environment, overlooking aspects such as the interpersonal characteristics of the supervisors and supervisees.

As mentioned earlier, Ellis et al. (2007) conducted a confirmatory factor analysis on data from the WAI-T and found that the three subscales were highly correlated. This led to the conclusion that the instrument may be measuring a single factor rather than three distinct factors. Other studies further suggest that Bordin's tri-factor model might be represented by a single factor (Andrusyna et al., 2001; Ladany et al., 1999). Furthermore, in its construction, the WAI was not subjected to factor analysis to test its factor structure with a sample of supervisees.
Thus, the concern that Bordin's theory might be best accounted for by a single factor structure could be a function of the untested WAI and warrants further study. In summary, it seems that the WAI-T may lack essential components to be used as a reliable measurement for supervision research.

**Supervisory Working Alliance Inventory (SWAI)**

The SWAI adapted elements from the body of literature on the working alliance in counseling (Gelso & Carter, 1985; Greenson, 1967; Patton, 1984; Pepinsky & Patton, 1971; Robinson, 1950) and the supervisory alliance model (Bordin, 1983). The items on the SWAI were developed using the test developers’ theoretically driven ideas and a task analysis of behaviors in supervision conducted by a group of expert supervisors. Designed to measure aspects of the relationship in supervision, the SWAI defined the working alliance as a set of actions interactively used by supervisors and supervisees to facilitate the learning of the supervisees (Efstation et al., 1990).

The SWAI items describe activities that represent target behaviors of supervisor and supervisee in supervision. The respondents on the SWAI are required to rate the degree to which target behaviors are performed in supervision. There are two forms of the SWAI: a 19-item supervisor's form and a 23-item supervisee's form. The original exploratory factor analysis yielded three orthogonal factors (i.e., client focus, rapport, and identification) for the supervisor's version and two orthogonal factors (i.e., rapport and client focus) for the supervisee's version (Efstation et al., 1990). Efstation et al. (1990) explained the differences between the number of factors on the two forms, suggesting that
supervisors have greater perception the complexity of the relationship due to their
greater knowledge and experience and, thus, the supervisor's version has more
factors then the supervisee's version. The SWAI has acceptable scale reliability;
the alpha coefficients for the supervisor scales are .71 for client focus, .73 for
rapport, and .77 for identification, and supervisee scales are .90 for rapport
and .77 for client focus. Correlations among the three supervisor's scales from the
SWAI are low but significant, ranging from .23 to .26. The correlation between
two supervisee's scales is .47. Also, correlations within each dyad between the
supervisor and supervisee scales range from .03 to .36.

Although, the SWAI has been used in a range of supervision related
studies (e.g., Humeidan, 2002; Sterner, 2009; Webb & Wheeler, 1998; White &
Queener, 2003; Wood, 2005), many researchers have criticized the psychometric
properties of the measurement. Ellis and Ladany (1997) pointed out that only
rapport on the supervisee scale has sufficient internal consistency ($\alpha = .90$) for a
scale measuring a single factor. Also within dyad correlations on the two versions
of SWAI, the scale ranged from non-significant to modestly significant (.03
to .36), thereby indicating a lack of agreement between two versions of the SWAI
on the construct of supervisory working alliance. One possible explanation is that
the SWAI is based only partially on Bordin's (1983) theory of working alliance,
thus the instrument may not correspond directly to the three elements of the
alliance as proposed by the model. Ellis and Ladany also suggested that the two
forms of the instrument may measure different constructs due to their having a
different number of factors.
Like the two previous instruments, the SWAI does not account for the developmental aspect of the supervision environment. For example, there are no items that address the supervisors' ability to understand and match their supervisees' developmental needs or account for the supervisors' ability to mismatch beyond the supervisee's current level of functioning. Instead, most of the items are focused on the skills and interventions that might enhance the supervisory relationship.

**Role Conflict and Role Ambiguity Inventory (RCRAI)**

The Role Conflict and Role Ambiguity Inventory (RCRAI; Olk & Friedlander, 1992) was developed to measure counselor trainees' experiences with role conflict and role ambiguity. One subscale measures role conflict that arises when individuals encounter opposing expectations for their behaviors. The second subscale measures role ambiguity that arises when individuals experiences a lack of clarity regarding their role. Twenty-nine items were constructed following content analysis of semi-structured interviews with 15 supervisors and trainees that asked about their experience of role difficulties within supervision.

Psychometric evaluation of the scale was conducted based on the responses of 240 doctoral-level counseling or clinical psychology trainees who completed the RCRAI as well as the Trainee Personal Reaction Scale (TPRS-R; Holloway & Wampold, 1984) measuring supervision satisfaction, the Job Descriptive Index (JDI; Smith, Kendall, & Hulin, 1969) measuring job satisfaction, and the State Trait Anxiety Inventory (STAI; Spielberger, Gorsuch, Lushene, Vagg, & Jacobs, 1983) measuring work related stress. A range of data
analyses (e.g., principal components analysis, scree test, and considerations of parsimony and interpretability) led to a 2-factor solution, which was supported by factor rotation. Internal consistency estimates using Cronbach’s alpha revealed coefficients of .91 for role ambiguity and .89 for role conflict. Convergent validity was demonstrated through the finding of a significant inverse relationship between the RCRAI scales and the measure of supervision satisfaction and job satisfaction. The results also demonstrated a significant positive relationship between RCRAI scores and the measure of work related stress.

Ellis and Ladany (1997) reported the RCRAI as a relatively sound measurement. The RCRAI was used in several supervision studies (e.g., Ladany & Friedlander, 1995; Nilsson & Duan, 2007; Nilsson & Anderson, 2004). However, this instrument, once again, fails to adequately address the developmental aspect of the supervision environment. Due to the purpose of this instrument, it measures perceived role conflict and role ambiguity in supervision, but not the underlying developmental problem of supervision that might affect role conflict and ambiguity. Furthermore, Ellis and Ladany pointed out that the strong correlation between the two scales as well as items loading highly on both scales may bring its factor structure into question.

**Supervisory Relationship Questionnaire (SRQ)**

The 67-item Supervisory Relationship Questionnaire (SQR; Palomo et al., 2010) was developed based on Beinart’s (2002) theory of the supervisory relationship. Using a grounded theory analysis, Beinart examined the factors that predict the quality of the supervisory relationship and identified nine factors that...
include: (a) boundaried, (b) supportive, (c) open, (d) respectful, (e) committed, (f) sensitive to needs, (g) collaborative, (h) educative, and (i) evaluative.

Principle component analysis of the SRQ revealed six components accounting for 65.3% of the variance (Palomo et al., 2010). Items loading onto single factors at greater than .4 were retained, which resulted in a 67-item scale. The six factors were used to create six subscales: (a) safe base; (b) structure; (c) commitment; (d) reflective education; (e) role model; and (f) formative feedback. The three subscales, safe base, commitment, and structure, reflect the facilitative relationship characteristic of the supervisory relationship. The other three subscales, reflective education, role model, and formative feedback, reflect the educative and evaluative functions of supervision. Among the factors, safe base accounted for most of the variance (52%) in the SRQ, which indicates its role as a powerful precondition for other aspects of the supervisory relationship.

The test developers suggested that the SRQ is a valid and reliable measurement (Palomo et al., 2010). The instrument showed high overall internal consistency ($\alpha = .98$), acceptable levels of convergent and divergent validity, and high test-retest reliability ($r = .97$). However, they also noted several limitations of the SRQ. First, the SRQ is a relatively new instrument that lacks additional validations. Ellis and Ladany (1997) recommended the use of confirmatory factor analysis in instrument development to validate the instrument model, thus the SRQ may benefit from a subsequent validation study using confirmatory factor analysis. Second, the norm sample was highly homogenous (e.g., mostly female clinical psychologist trainees); therefore administering the SRQ to other
populations to determine its generalizability is recommended. Additionally, the SRQ has not been widely used in current supervision research or practice, and no subsequent publications from its authors have been presented in the professional literature. Similar to the previously mentioned measurements, the SRQ does not appear to adequately address the necessary elements of the optimal supervision environment. For example, the SRQ was not developed to measure the supervisors' ability to provide adequate support and challenge. Furthermore, very high overall internal consistency may indicate that the scale is measuring a narrow construct range, while supervision is highly complex. High test-retest reliability also may indicate that the measure is not sensitive to change; thus, this instrument may not necessarily incorporate the concept of change over time in considering the needs and responses of supervisees.

Need for a New Measure

In summary, a careful review and evaluation of existing supervision models and instruments for content and psychometric soundness revealed that they were inadequate to fully address the components of the optimal supervision environment. Some instruments suffered from less than rigorous methodology in terms of instrument design, item development, and validation procedures. Ellis and Ladany (1997) recommended the use of confirmatory factor analysis in instrument development and the development of an a priori factor model in supervision; however, no supervision measurement has used such analysis. Furthermore, existing instruments do not account for the developmental aspect of the supervision environment. Therefore, it seems that in order for supervision
research to advance, a new valid and reliable instrument that measures the supervisor's ability to create an optimal supervision environment is warranted. The following sections will describe the development of a measurement of the optimal supervision environment using attachment theory (Bowlby, 1988) and constructive developmental theory (Kegan, 1982). These two theories are proposed as robust frameworks for understanding the complexity of the supervision environment.

Theoretical Considerations

Little research has been conducted with regard to assessing the optimal supervision environment. The current study examined the optimal supervision environment from two developmental theories: constructive developmental theory (Kegan, 1982, 1994) and adult attachment theory (Bowlby, 1998). These theories are thought to provide rich, solid frameworks to understand and identify the elements of the optimal supervision environment. They appear to incorporate individual characteristics of the supervisor and supervisee as well as differences in the psychological capacities of adult learners that impact the dynamics of the supervision environment. Specifically, the two theories will serve as the foundation for developing of a new scale designed to measure the optimal supervision environment.

Constructive Developmental Theory

Piaget (1970) posited that universal structures of knowing shape the framework by which people view the world. The structures both define and limit one's view of the world, including self-awareness and perceptions of others.
Piaget argued that individuals grow developmentally through the interaction with their environment. When new information is encountered the individual first attempts to comprehend it in terms of what is already known (assimilation). When the new information is too divergent to fit existing structures, disequilibrium occurs, and the insufficient mental structure must be replaced with a structure that is a better fit for comprehending the new environmental information (accommodation). Piaget’s research laid the foundation for numerous developmental theories, including Perry’s (1970) work on intellectual development, Kohlberg’s (1981) work on moral development, Selman’s (1980) work on social perspective taking, Loevinger’s (1976) work on ego development, and Kegan’s (1982, 1994) constructive-developmental theory.

Based on Piaget’s (1970) work, Kegan (1982) posited a theory of development, referred to as constructive-developmental theory (CT). CT is premised on two fundamental concepts: (a) a constructivist perspective and (b) a developmental perspective. A constructivist perspective proposes that individuals are continually engaged in the active process of constructing their reality. Kegan (1982) stated that the activity of being a person is the activity of meaning-making. There is no feeling, no experience, no thought, and no perception independent of a meaning-making context. The way in which individuals experience the world is dependent upon how they mentally organize it.

A developmental perspective proposes that mental structures evolve through qualitatively different periods of growth, based upon alternating periods of stability and change (Kegan, 1982). CT suggests that the mental structure
individuals use to organize meaning-making changes and evolves systematically and in stages. Individuals develop as they gradually incorporate their current mental structure into a more complex mental structure, and their environment is the key that supports or constrains the process of development. CT integrates the constructivist and developmental concepts into a theory on development of meaning-making (i.e., constructing reality) and evolving consciousness that extend Piaget’s model of stages of development into adulthood (Kegan, 1982). Kegan (1982) identified the structures of knowing underlying the development of individual’s self conception and capacity for relating to others, and referred to the structural stages as the Subject-Object balance.

**Subject-Object Balance.** Kegan (1984, 1992) explained the individual development process in structural or subject-object terms. Using the concepts of Object Relations theory, he asserted that individuals define self through a process of referring to self in relation to others rather than defining self as an isolated entity. “Subject” refers to the organizing principle of reality by which one knows (Rogers & Kegan, 1991, p. 105). Individuals are “identified with, tied to, fused with, or embedded in” those aspects of the environment and self that are “subject” (Kegan, 1994, p. 32). “Objects,” on the other hand, are those aspects of experience that people can “reflect on, handle, look at, be responsible for, relate to each other, take control of, internalize, assimilate, or otherwise operate on” (Kegan, 1994, p. 32).

“Subject” becomes “object” as individuals develop to the next order of consciousness. Kegan (1982, 1994) outlined six stages of consciousness.
development across the lifespan: the incorporative, the impulsive, the imperial, the interpersonal, the institutional, and the interindividual. Infants in the incorporative stage are subject to their reflexes, and they have no object on which they can reflect (stage 0). Children at the impulsive order of consciousness (stage 1) are aware of their reflexes as objects but they are subject to their impulses. Children and teenagers who are at the imperial order of consciousness (stage 2) can reflect on their impulses and perceptions, but they are subject to their individual needs, interests, and desires. Teens and young adults in the interpersonal order of consciousness (stage 3) are able to reflect on their individual needs, interests, and desires but they are subject to interpersonal relationships and mutuality. Adults at the institutional order of consciousness (stage 4) are able to reflect on interpersonal relationships and mutuality, but they are subject to their own authorship, identity, and ideology. Finally, those in the interindividual order of consciousness (stage 5) are able to reflect on their own authorship, identity, and ideology while being subject to “the interpenetrability of self-systems” (Kegan, 1982, p. 82).

One of the assumptions of Kegan’s stage theory is that meaning occurs in evolving stages of becoming less rigid, simplistic and dogmatic and more flexible, open, complex, empathic, and tolerant of difference. However, the assumption that higher developmental stages are better does not suggest that one at a higher stage is superior in all areas, but, rather, it suggests that one has “better conceptual tools for making sense of the world and deriving guides for decision making” (Rest & Narvaez, 1994, p. 17). Promoting development becomes particularly
important when one considers that most individuals lack opportunities for the type of interactions that stimulate growth (Manners, Durkin & Nesdale, 2004). Development occurs when individuals are faced with challenging new experiences that create discomfort or disequilibrium. However, these challenges alone are not sufficient, and development also requires a supportive environment. This contention has important implications for understanding the complexity of counselor development and the role of the supervision environment.

Kegan (1984, 1992) asserted that people need an environment that concurrently attends to the stage from which they are functioning and to which they are transitioning. The key to a helpful environment is that it must match (support) people by relating to them from their currently dominant way of knowing and mismatch (challenge) people by relating to them from the next potential way of knowing. Morgan, Morgan, Foster, and Kolbert (2000) stated that development does not occur automatically, but must be stimulated given an adequate learning environment that includes opportunities for role-taking, support, challenge and guided reflection. In addition, Border (1998) encouraged incorporation of “optimal environment” that provides an appropriate balance of challenge and support, innovation, and integration. Given such an environment, supervisees feel secure enough to take risks and are better able to achieve insights characteristic of the next developmental level.

The Holding Environment. Similar to Border’s (1998) emphasis on the need for an “optimal environment,” Kegan (1982) also called attention to the importance of what he calls a “holding environment” (p. 116). Winnicott (1965)
first used the term “holding environment” to refer to the psychosocial environment that supports the healthy development of an infant. Development is affected by the unique interaction of individual needs and strengths and the particular situational forces or holding environments in which we are situated (Winnicott, 1965). Kegan noted the importance of infant and early childhood holding environments and expanded the idea to suggest that new holding environments that come late in life may also contribute significantly to the development of the self. He characterized the holding environment as “the social, physical, psychological context(s) in which and through which an individual develops and comes to know and define his very self” (Kegan, 1982, p. 52). He described the concept of holding as “not to keeping or confining, but to supporting the exercises of who the person is” (Kegan, p. 162). He suggested that the self becomes redefined as a child attaches, pulls away from and reintegrates a new sense of self apart from each holding environment. A healthy holding environment can affirm individuals as they are as well as assist in their development.

A good holding environment carries out three major functions: holding on, letting go, and maintaining (Kegan, 1982, 1994). First, the environment must hold well, meaning that it recognizes and confirms individuals as they are currently making meaning without creating frustration or demands for change. Second, a good holding environment lets individuals move on when they are ready, challenging them to grow beyond their existing perceptions to new and more complex ways of knowing. The holding environment needs to challenge the
learners to question and rethink their constructions of self and ways of knowing at a particular time (Kegan et al., 2001). Third, a good holding environment remains in place to recognize and sustain individuals' growth and change. It provides continuity, stability, and availability to the person in the process of growth.

The holding environment encourages growth when it supplies an optimal balance of challenge, support, and continuity, according to the specific requirements of one's stage of development (Kegan, 1982, 1994). Too much support without enough challenge may be comfortable but insufficiently stimulating, and individuals that are overly supported may feel bored or disengaged. Conversely, too much challenge without enough support can generate defensive resistance and withdrawal, and individuals who are overly challenged may feel threatened, alienated, and overwhelmed.

In summary, Kegan's (1982, 1994) theory offers a new understanding of adult development which has the potential for illuminating the essential components of counselor development. Kegan, in an interview with Eriksen (2006), spoke of creating a supervisory relationship that serves as a bridge from where supervisees should be able to stand at any point on that bridge and feel well supported. This approach seems to be supported by Worthen and McNeill's (1996) finding that good supervisors are seen by their supervisees as empathic, nonjudgmental, validating, non-defensive, and willing to examine their own assumptions in order to normalize their supervisees' struggles while encouraging them to explore and take risks.
A significant body of counseling research has supported the claim that counselor trainees with higher developmental levels are more capable of several of the tasks of counseling such as increased empathy, less negative bias, more autonomy, more flexibility, and better problem solving when working with diverse individuals (Foster & McAdams, 1998; Lambie, Ieva, & Mullen, 2011). For example, Borders (1989a), as mentioned earlier, explored the relationship between ego developmental levels and their perceptions of clients and in-session cognitions. Results indicated that students with higher ego levels had significantly fewer negative thoughts about their clients, were less critical of themselves, and possessed more objective retrospections overall. In addition, high cognitive complexity levels have been associated with increased ability to empathize with others (Perry, 1970) and, in particular, the ability to appropriately empathize with clients from different cultures (Chung & Bemak, 2002; Granello, 2002; Frame & Williams, 2005). Finally, levels of cognitive and emotional development may influence the depth at which counselor trainees conceptualize client problems, process the therapeutic relationship, recognize affective changes, deal with a manipulative client, and handle termination (Holloway & Wampold, 1986; Loganbill, Hardy, & Delworth, 1982).

In accordance with developmental theory, a supervisee’s way of knowing can become more complex if he or she is provided with developmentally appropriate supports and challenges. Attending to the diversity of ways in which supervisees interpret and make sense of their experience can provide new and important insights into their counseling and supervision experiences. Because
development is relational, taking place in the social context, supervision should provide an optimal environment for continued development (Hayes, 1994). A more complex understanding of counselor development through the Kegan’s concept of holding environment could enhance the capacity of supervisors and counselor educators to tailor their interventions to the developmental needs of the trainee. In addition, Kegan’s constructive developmental theory could provide a framework for using diverse approaches in meeting and supporting learners with a diversity of learning needs and ways of learning.

Attachment Theory

Bowlby (1988) proposed attachment theory to understand how certain early emotional bond experiences influence emotional and physical well-being, not only in childhood but throughout adulthood as well. Bowlby proposed that each individual develops an attachment behavior system which is an inborn regulation system that has important implications for personality development and social behaviors. This attachment system is activated by perceived threats and dangers that cause a threatened person to seek proximity to a caregiver (Mikulincer & Shaver, 2007). However, when the quality of the relationship with the caregiver is inadequate, the relationship can in fact be a major source of stress (Repetti, Taylor, & Seeman, 2002), and when attachment security is not achieved, the use of alternative, insecure attachment strategies of avoidance or anxiety are triggered (Ainsworth, Blehar, Waters, & Wall, 1978). Because healthy functioning of the attachment system facilitates relaxed and confident engagement in nonattachment activities, it also contributes to the broadening of individuals’
perspectives and skills, as well as the actualization of their unique potential (Mikulincer & Shaver, 2007).

Attachment behavior is defined as any activity that results in proximity or an enduring emotional bond to a preferred other (i.e., a caregiver) who is considered as stronger and wiser. For example, Simpson, Rholes, and Nelligan (1992) reported that secure females whose behavior indicated anxiety about an impending stressor sought contact with their partners, presumably to reduce anxiety. Similarly, Fraley and Shaver (1998) observed and coded the behaviors of couples awaiting a separation (which was assumed to be anxiety provoking) in an airport to determine how attachment behaviors are manifested during separations and how the organization of attachment behavior is impacted by factors known to regulate attachment behaviors in infancy (i.e., accessibility of the attachment figure, length of the relationship, and working models). The researchers reported that proximity maintenance also appears to be a function of separation behaviors in adults. For example, adults, who were about to experience a separation, were likely to hold onto, follow, and search for their partners. Their observed behaviors were considered as efforts to modulate anxiety.

Researchers have identified several models of attachment styles that reflect differences in the working models of self and other (Bartholomew, 1990; Bowlby, 1973). Attachment styles are generally referred as different categories of attachment quality including patterns of beliefs, needs, emotions, and social behaviors that result from particular attachment experiences (Fraley & Shaver, 2000). Attachment styles were first identified in an infant observational study
based on Bowlby’s theory (Ainsworth et al., 1978). According to attachment theory, different patterns of attachment emerge in response to the way caregivers react to their infants’ attachment behaviors (Ainsworth et al., 1978). Three major attachment styles were identified: (a) secure attachment style, characterized by comfort in getting close and depending on others, (b) avoidant attachment style, characterized by difficulty getting close and depending on others, and (c) anxious-ambivalent attachment style, characterized by ambivalence towards caregivers. Infants tend to develop secure attachments to caregivers who serve as good, receptive targets for their attachment behaviors. Caregivers who foster greater security tend to read their infants’ cues of distress more accurately and find effective ways to comfort them. When their infants are not distressed, the caregivers remain physically and emotionally available to their infants without being disruptive or intrusive. Infants who form secure patterns of attachment with their attachment figures differ from insecure infants in numerous ways.

The Strange Situation test (Ainsworth et al., 1978) involved a series of distressing separations and reunions between mothers and their 12- to 18-month-old children. The results of the test revealed that those children classified as having secure relationships directly searched for comfort from their mothers, were calmed easily, and then resumed other activities (e.g., playing, exploring room). Children classified as having anxious-ambivalent relationships displayed decidedly mixed reactions to their mother (i.e., approach-avoidance behaviors), remained agitated, and failed to resume normal activities. Children classified as having avoidant attachment disregarded their mothers, showed signs of emotional
disengagement and withdrawal, and engaged in behaviors that kept them
distracted from the distress they were feeling. In subsequent research, Main and
Solomon (1990) reanalyzed Ainsworth and her colleagues' data and identified a
fourth category, the disorganized-disoriented attached style, characterized by
contradictory and confused behaviors toward caregivers.

Attachment relationships can lead to the development of internal working
models that allow for the prediction of an individual's future proximity-seeking
behaviors with significant others (Main, 1996). These internal working models
provide "rough-and-ready blueprints" (Rholes & Simpson, 2004, p. 7) for what
should be expected and what is likely to occur in different kinds of interactions
with attachment figures. Working models organize affect, behavior, and cognition
in close relationships, providing guidance about how to behave, what should be
expected or anticipated, and how to interpret the meaning of ambiguous
interpersonal events. Working models are termed working because they remain
open to correction and revision. Similar to schemas in developmental theories,
working models tend to be conservative, in that new experiences are assimilated
into existing models more readily than models are accommodated to fit new
experiences. Therefore, working models are situated at the juncture of existing
premises and new information.

In particular, the attachment and caregiving behavioral systems are
important to relationship functioning and to the security that supports optimal
development. The early organization of individuals' attachment behavior
functions throughout life by motivating them to gain security (Bowlby, 1988).
The individuals gain security by maintaining proximity to a caregiver, which, in turn, shapes the way they see the world about them and the manner in which they expect persons to whom they might become attached to behave. As an adult, the internal working models become central features of personality that direct cognitive, affective, and behavioral components of attachment behavior (Bowlby, 1988; Main, 1990). Additionally, relatively stable working models are open to revision and become increasingly complex as individuals have new interpersonal experiences and are influenced by significant attachment-related experiences such as psychotherapy (Bowlby, 1988) and marriage (Crowell, Treboux, & Waters, 2002).

**Adult Attachment Theory.** Bowlby (1988) suggested that the fundamental tenets of attachment occur throughout life in child-parent and significant adult relationships. Adult attachment involves a dyadic relationship in which proximity to a significant other is sought or maintained to achieve a sense of security (Sable, 2008). According to adult attachment theory, even securely attached adults seek relational proximity to others in order to promote, enhance, or restore a perceived sense of security. Early attachment experiences are carried forward as mental representations of attachment figures in relation to the self along the pathway toward developing the adult personality (Sable, 2008).

Bartholomew and Horowitz (1991) proposed a model for the internal working models of adult attachment comprised of two dimensions: view of *self* (positive or negative) and view of *others* (positive or negative). This model identified four prototypical adult attachment styles based on the internal working
models of *self* and *others*: secure (positive view of self, positive view of others), preoccupied (negative view of self, positive view of others), dismissive (positive view of self, negative view of others), and fearful (negative view of self, negative view of others).

According to attachment theory, the caregiver provides a secure and dependable base for the child to explore the world referred as the secure base. This secure base impacts individual’s significant relationships (e.g., child-parent relationships, student-teacher relationships, and romantic relationships) by determining how they as infants, children, or adults, behave with the outside world (Newwald-McCalip, 2001). For example, secure children display confidence exploring away from the immediate proximity of their primary caregiver, knowing the caregiver will be accessible upon return. As the children become older, their exploration expands, but the secure base remains necessary for optimal functioning and mental health (Bowlby, 1988). In adulthood, the existence of a secure base continues to influence whether individuals will exhibit confidence in seeking out help when needed or in exploring diverse new roles and settings (Newwald-McCalip, 2001).

Green and Campbell (2000) examined the link between attachment and exploration. They looked at 100 undergraduate students (79 females, 20 males, 1 not reported) who completed an exploration index and attachment checklist (Simpson et al., 1992). The researchers developed the exploration index on which students rated their interest and likelihood of engaging in a variety of novel activities. Results indicated that both anxiety and avoidance attachment style
correlated with exploration, with avoidance generally yielding a stronger association. In particular, avoidance highly correlated with reduced social exploration, and anxiety highly correlated with reduced environmental exploration. The results were evaluated in relation to a measure of attachment styles and found that security predicted higher exploration scores. The study provides direct evidence that attachment is associated with orientation toward exploration. The researchers recommended future studies that may further illuminate more specific relationships between the two insecure styles and different types of exploration.

**Attachment Theory and the Supervisory Relationship.** Bowlby (1988) claimed that in psychotherapy the therapist should assume the role of an attachment figure, providing a secure base from which clients may confidently explore and reassess their working models of attachment figures and of themselves. Furthermore, Watkins (1995) suggested that the supervisory relationship also has many similarities to both parent-child relationships and many adult-adult relationships. Internal working models are dynamic representations that may be altered in response to new role taking experience. In other words, supervisees’ working models of self and others can be modified through ongoing interpersonal relationships, increased self-understanding, and renegotiations of the balance between connection with others and independence (Neswald-McCalip, 2001). Once grounded in a secure relationship, supervisees are freed to experiment with particular techniques, becoming more creative in session, or consulting with other professionals and colleagues (Pistole & Watkins, 1995).
Therefore, because relationships between supervisors and supervisees are often characterized by essential elements of attachment relationships (Pistole & Watkins, 1995), a secure supervision base can serve as a modifier for the supervisee's current attachment style. Worthington and Roehlke, (1979) reported that supervisors who encouraged their supervisees to try new therapeutic approaches and develop an independent counseling style were likely to be available during crises and offer reassurance, qualities similar to the secure base and safe haven aspects of an attachment relationship.

Fitch et al. (2010) developed a supervision model called the Attachment-Caregiving Model of Supervision (ACMS) that applied the attachment behavior system theory to the supervisory relationship. The ACMS examines the caregiving and attachment processes in the supervisory relationship and their link to learning. Specifically, ACMS addresses the conditions that facilitate bonding, the quality of supervisors and trainee bonds, the mechanisms (e.g., affect regulation, sensitivity to attachment cues) through which the relationship is co-constructed and functions, and how the constructed relationship and its functions relate to learning. The authors claimed that supervisors provide an attachment safe haven function by being attentive and alert to attachment cues and deciphering and responding to those cues in a way that corresponds to the supervisee's working model, thus alleviating their anxiety by maintaining appropriate proximity and safety. For a beginning supervisee, the supervisor may provide the safe haven by combining structure and support with didactic feedback;
whereas with a more advanced supervisee, the supervisor may focus on clarifying his or her comprehension and perspective.

Conceptualizing the supervisory relationship as one in which the supervisor functions as a secure base and a haven of safety suggests the need for specific strategies that supervisors can implement to maximize their effectiveness in performing that function. When supervisors are able to facilitate conditions under which secure attachment develops, supervisees are likely to be training in an optimal supervision environment.

Several studies demonstrated important implications of attachment theory with regard to the supervisory relationship and supervision environment. Foster et al. (2007) conceptualized the supervisory relationship as one in which the supervisor functions as a secure base and a haven of safety. They used 90 supervisor-supervisee dyads to examine attachment style and professional development. The Supervisee Levels Questionnaire-Revised (SLQ-R; McNeill et al., 1992) was used to measure supervisee self-reported professional development, the Supervisee Levels Scale (SLS; Wiley & Ray, 1986) to measure supervisee professional development rated by the supervisor, and the Relationship Scales Questionnaire (RSQ; Griffin & Bartholomew, 1994) to measure attachment. The researchers found that supervisees who were insecurely attached to their supervisor reported lower levels of professional development compared with trainees who were securely attached to their supervisor. The study noted some limitations such as the relatively small effect size of the relationship between attachment and professional development and lack of psychometric validation for
the use of RSQ to evaluate supervisory relationship. The authors recommended the examination of the impact of supervisors' own attachment on the quality of the supervisory relationship, acknowledging that the supervisor's level of attachment security likely plays an important role in the quality of supervisory relationship.

White and Queener (2003), as mentioned earlier, found that the supervisor's ability to make attachments and social provisions (i.e., social network) are more predictive of the quality of the supervisory alliance than are the same characteristics that the supervisee brings to the relationship. One explanation they pointed out for the finding is the fact that the supervisory relationship is a structured, hierarchical relationship in which the supervisor is more powerful; thus, the characteristics of the supervisee had a lesser impact on the supervisory relationship. The authors pointed out that most models of supervision do not explicitly consider the individual characteristics of the supervisor and supervisee in understanding and explaining the dynamics of the supervisory relationship.

Riggs and Bretz (2006) conducted a study looking at how attachment styles of clinical trainees and supervisors influence the supervisory relationship. They surveyed 87 doctoral-level psychology interns about their attachment processes and supervision experiences. The Working Alliance Inventory (Horvath & Greenberg, 1989) was used to assess the supervisory working alliance. The Measure of Parental Style (Parker et al., 1997) was used to assess early parent-child relationships. The Reciprocal Attachment Questionnaire (West &
Sheldon-Keller, 1994) was used to assess pathological attachment behaviors. Participants also reported their own attachment style and their perceptions of their supervisors' attachment styles using the Relationship Questionnaire (RQ; Bartholomew & Horowitz, 1991). Results showed that the attachment style perceived in the supervisor was significantly associated with the supervisory task and bond. Regardless of supervisees' perceived attachment style for themselves, those who perceived their supervisors to be securely attached rated the supervisory bond higher than those who perceived their supervisors to be insecurely attached. In other words, perceived supervisor attachment styles were important elements impacting on the supervisory working alliance. Like White and Queener (2003), Riggs and Bretz (2006) argued that the hierarchical nature of the supervisory relationship, in which the supervisor has greater power and greater experience, suggests that the supervisory working alliance is more likely to be influenced by supervisor attachment style than by supervisees' attachment style.

Dickson et al. (2011) replicated the study of Riggs and Bretz (2006) but examined the elements affecting the supervisory alliance with a sample of British clinical doctoral supervisees that was larger and more homogeneous. As with other studies, Dickson et al. found that supervisees' ratings of the supervisory working alliance were associated with perceptions of their supervisor's attachment style, but not with their own attachment style. Ratings of the supervisory working alliance were lower when trainees perceived their supervisors to be insecurely attached, irrespective of trainees' self-reported
attachment style. Despite some limitations of this study (e.g., biases from self-reporting, limited range of sampling, and absence of supervisors’ perspective), the findings further indicate that supervisor attachment style may have a significant impact on the supervisory environment.

**Purpose of the Study**

The present study justifies and presents the development of a new instrument that measures the quality of the supervision environment to be referred to hereafter as the Chae Optimal Supervision Environment Test (COSET). The previous review of current supervision models and instruments revealed a crucial need for a valid, reliable instrument that assesses the quality of the supervision environment as a venue for promoting counselor development. It is intended that the COSET could serve as a new method of measuring the supervisors’ ability to create a “good enough” supervision environment in order to promote counselor and supervisor development. Given the lack of empirically validated measurement tools in the area of supervision in general (Ellis & Ladany, 1997), this new instrument could provide a foundation for further studies in supervision. The COSET is expected to be applied across the broad fields of study, including counseling, supervision, counselor development, supervisor development, and teacher development.

**Research Objectives**

**Research Objective 1**

To review the related literature to identify, describe, and define potential elements of an optimal supervision environment.
Research Objective 2

To create the preliminary items of the COSET.

Research Objective 3

To test the initial reliability of the COSET.

Research Objective 4

To examine the correlations among factors of the COSET.

Research Objective 5

To identify the factor structure for the COSET.

Conclusion

This chapter described the relevant problems in clinical supervision (i.e., negative supervision experiences, lack of supervision research, and lack of reliable, valid supervision measurements) and the current approaches to those problems. The IDM model and the working alliance model were described as providing various benefits to the counseling and supervision fields; however they were each shown to be limited in comprehensively capturing the complexities of the supervisory relationship. Instruments for measuring supervision quality were examined and found to be of limited utility due to less than rigorous methodology in their development. The current supervision literature clearly points to the need to develop reliable and valid means of supervision quality assessment. The literature also affirms the promising potential of constructive developmental theory and attachment theory as useful frameworks to identify interpersonal and developmental elements in the supervision environment that affect the supervisory relationship and supervision outcomes.
CHAPTER TWO

METHODS

This chapter describes the research methodology for a scale development study, including the instrument construction, field-test sample, and data collection procedures that underpin the development of the Chae Optimal Supervision Environment Test (COSET).

Procedures for Construction of the COSET

Item Development

The procedures detailed in this chapter describe the COSET scale development process from construct conceptualization through item development and field-test administration. Construction and field-testing of the COSET took place in five phases. The first phase involved reviewing the relevant literature to identify important elements of supervision to serve as the basis for creating a blueprint for the COSET. The scale blueprint delineated three scales (i.e., Emotional Environment, Learning Environment, and Power Environment); each deemed to be separate but important aspects of counselor supervision. The second phase involved writing items to populate the test blueprint cells, along with multiple iterations of item editing and modification. The third phase involved an empirical evaluation of the items by piloting the COSET in an effort to improve the clarity of items and reduce their total number. The fourth phase included a review of items by a panel of supervision experts. The final phase involved a larger scale field-test study that included administering the COSET to a national sample of clinical supervisors and counselor educators to provide data.
for analyses that would lead to further refinement of the instrument, as well as to collect evidence for the investigation of the scale’s reliability, validity, and factor structure.

First Phase: Preliminary Item Development. To develop the COSET, the researcher, first, carefully reviewed and evaluated existing supervision models and instruments for content and psychometric soundness. The researcher, then, conducted a comprehensive review of literature in the area of counseling supervision to identify the essential elements of an optimal supervision environment. Constructive developmental theory and adult attachment theory provided the frameworks for identifying the conceptual elements. The literature was drawn from the fields of counseling, counselor education, and clinical supervision. To provide the broadest search possible, research articles, dissertations, and books were accessed. Only works published after 1995 were included in the review to provide a current conceptualization of clinical supervision. From this literature review, the researcher identified limitations of clinical supervision and supervision research and instrumentation, and developed a blueprint for the COSET to assess supervisors’ perceptions of these salient environmental issues. The researcher defined the construct (i.e., optimal supervision environment) and distilled from the literature three essential components of an optimal supervision environment. These three elements, repeatedly identified and examined by other researchers, included the: (a) Emotional Environment, (b) Learning Environment, and (c) Power Environment (e.g., Barrett & Barber, 2005; Borders, 1998; Dickson et al., 2011; Ellis, 2001;
Falender et al., 2004; Fitch et al., 2010; Foster et al., 2007; Holloway, & Neufeldt, 1995; Kilminster & Jolly, 2000; Ladany et al., 1999; Ladany, Lehrman-Waterman, Molinaro, & Wolgast, 1999; Magnuson et al., 2000; Neswald-McCalip, 2001; Palomo et al., 2010; Riggs & Bretz, 2006; Stoltenberg, 2005; Wheeler & Richards, 2007; White & Queener, 2003; Worthen & McNeill, 1996). Additionally, Kegan’s (1984) three components of the *holding environment* (i.e., support, challenge, and continuity/maintainability) served as the structure for each environment during the item development. Figure 2.1 depicts the COSET blueprint. The COSET blueprint was constructed to have equal numbers of items reflecting the three elements, but the three components of the holding environment were only guiding elements and were not represented equally.

Figure 2.1

**COSET Blueprint**

<table>
<thead>
<tr>
<th>Elements of COSET</th>
<th>Emotional Environment</th>
<th>Learning Environment</th>
<th>Power Environment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Support</td>
<td>Support</td>
<td>Support</td>
<td></td>
</tr>
<tr>
<td>Components of Holding Environment</td>
<td>Challenge</td>
<td>Challenge</td>
<td>Challenge</td>
</tr>
<tr>
<td>Continuity/ Maintainability</td>
<td>Continuity/ Maintainability</td>
<td>Continuity/ Maintainability</td>
<td></td>
</tr>
</tbody>
</table>

*Second Phase: Item Development.* The researcher, then, wrote approximately 300 items for the COSET that collectively reflected the three
dimensions of supervision environment extracted from the literature review. The researcher wrote more items than necessary to over-sample each domain and ensure that the items broadly revealed as many facets of the domain as possible. The items were intentionally worded as concisely and briefly as possible.

The COSET was designed as an attitudinal measure that uses a Likert scale response format. Likert scales are routinely used with attitude or opinion scales and are well-suited for statistical tests of interval data (Cohen, Swerdlik, & Sturman, 2012). Likert scales also are considered ideal for measuring opinions, beliefs, and attitudes, which was consistent with the intended use of the COSET. To avoid response sets of central tendency, the items were constructed using 4-point response options, with no neutral option. The Likert scale response options were: 1 = strongly disagree; 2 = disagree; 3 = agree; 4 = strongly agree. After the items were written, an expert in test construction and psychological assessment with numerous published assessments edited the items. Following content review, the initial item pool was reduced to 200 items, which were tried out in the pilot study. The researcher used a web-based survey venue, Qualtrics, to create the survey and conduct the pilot and later field studies.

Third phase: Pilot Study. During the third phase of the study, the researcher conducted a pilot field study of the COSET with 14 doctoral students and faculty members from the researcher’s counselor education program. The pilot study was conducted to assess the instrument’s clarity, conciseness, readability, distinctiveness, and content reflection. All participants in the pilot study had experience as a supervisor, having provided supervision to one or more
professionals for at least one year. Each reviewer received an email consisting of a cover letter and a link to the preliminary COSET items. The willing participants reviewed the 200 item version of the COSET after being asked to reflect on one supervisee from a previous supervision experience, and then complete the COSET. The supervisors also were asked to identify any perceived problems with the questionnaire.

After collection of 14 cases, the preliminary items were analyzed using SPSS for Window 20.0 to conduct descriptive statistics and determine the initial scale’s reliability. Item-descriptive statistics (i.e., response frequencies, means, standard deviations, and range) were conducted to identify and modify items that were difficult to answer, and to delete items that did not contribute to the instrument’s variability. Internal consistency reliability for each scale was calculated using Cronbach’s alpha. As a general rule of thumb, measures of internal consistency are preferred to test-retest reliability as indexes of reliability (Wasserman & Bracken, 2013). Wasserman and Bracken (2013) suggested that scales intended for research applications should minimally be reliable at a level of .70, and preferably .80. Thus, for the purpose of this study, the reliability criterion is set at a level of .80. Cronbach’s alpha coefficient is a widely used method for computing the internal consistency form of reliability (Gall et al., 2007). Nunnally and Bernstein (1994) stated that “Coefficient α usually provides a good estimate of reliability because sampling of content is usually the major source of measurement error for static constructs” (p. 252). Cronbach’s Alpha “If Item Deleted” statistics was used to modify or remove items that were not
performing well, thereby increasing the reliability within each scale. Also, from the participants’ comments, the researcher modified or deleted items that lacked clarity, conciseness, readability, distinctiveness, and content reflection. The researcher modified the instrument by reducing or editing items, resulting in a total of 81 items retained at this phase, including 27 items per factor. The preliminary reliability for the total scale was .95.

**Fourth phase: Review of Expert Panel.** An expert panel of supervisors was used to conduct an initial assessment of COSET face and content validity. Five individuals comprising the expert panel reviewed the retained COSET item stems for readability and clarity. The raters are considered to be experts in the field based on their extensive research experience, scholarly research in supervision, and experience in providing supervision. Expert review is a critical procedure of scale development because it helps support the content validity of the measurement (Worthington & Whittaker, 2006). The raters were provided with a brief review of the literature detailing the characteristics of the optimal supervision environment and a form that contained all of the preliminary items of COSET. The researcher asked the expert panel to review the COSET supervisory elements and related items, and to examine the relevance of the items within each respective supervision environment. The reviewers rated the items based on three response options, including level of agreement that the item was suitable as written. The reviewers had three options regarding their dispositions toward the items and their fit to the COSET model and overall quality, including: agree (1), disagree (2), or modify (3). If the experts disagreed with the item fit and quality
or commented on modifying an item, they were asked to document their rationale for elimination or modification. The criterion for item determination was the support of at least three of the five experts who agreed to either to add, remove, or modify items. Raters were also asked whether they thought any important aspects of the supervision environment were missing from the item content. Based on the raters’ responses and comments, the researcher modified and eliminated several items, resulting in a total item pool of 78 items, with 26 items per subscale.

**Fifth phase: Administration of Initial COSET.** The purpose of the fifth phase of the study was to: (a) collect data on the COSET for item analyses and reduction; (b) provide evidence of construct validity; (c) examine the instrument’s factor structure; and (d) calculate final reliabilities for the total scale score and each of the three subscales. A national sample of 93 participants contributed to the data upon which analyses were conducted. The national sample is described in detail in the Participants section of this report.

Descriptive statistics were used to characterize the demographic background of the participants, as well as to analyze the means and standard deviations of scale responses. Prior to conducting analyses, the data were inspected to determine if there were any systematic response patterns, missing responses, or other anomalies. As a result of this quality control effort, 15 cases with missing data and 7 cases with irregular response patterns were eliminated. Moreover, given a four point response scale, eight items were eliminated due to ceiling effects; that is, they had both high means (M > 3.5) and low variability (SD < .5). These items were discarded because they failed to discriminate
between levels of supervision quality. The remaining items were analyzed to
determine scale reliability, a multiple iteration process that eliminated one item at
a time in an effort to contribute to scale refinement and item reduction. As in the
earlier analyses of the pilot data, Cronbach’s alpha coefficient (criterion > .80;
Wasserman & Bracken, 2013) was used to evaluate the internal consistency
among both items in the COSET total scale and for each of the subscales.

Corrected Item-Total Correlation and Cronbach’s Alpha “If Item Deleted”
statistics were used to remove items while sustaining high reliability within each
scale; that is, both statistics forecast the improvement or decrement to the scale
reliability if the item were to be removed from the instrument.

In addition, a series of exploratory factor analyses (EFA) were conducted
using a principle components analysis (PCA) with a Varimax rotation to estimate
the total variance explained by the specific items, to reduce the data set into a
smaller number of variables, and to reveal the underlying structure of the COSET.
The eigenvalues and Cattell scree plot were used as criteria for factor retention in
his study. Eigenvalues less than 1.0 reflect potentially unreliable factors (Kaiser,
1958); therefore, scales were retained if they evidenced eigenvalues equal to or
greater than 1.0. The Cattell scree plot and eigenvalues are generally used in
combination to determine the numbers of factors to be retained. The criteria for
factor retention were: (a) each factor had the eigenvalues equal to or above 1.0
from the correlation matrices; (b) a break in the shape of the Cattell scree plot at
the point of change in the elbow, and (c) a minimum item loading of .40. Any
item that loaded less than .40 was excluded from the instrument. The researcher also removed items that loaded more than .40 on multiple scales.

The instrument was then modified based on multiple iterations of reliability analysis, consideration of content sampling contributions, and factor analysis as mentioned above, with the goal of reducing the number of items to an optimal level. Ideally, a research-based instrument should have a short administration time, while maintaining high reliability and equal numbers of items across the three scales to maintain comparable scale contributions to the total test score (i.e., different numbers of items across scales would result in differential scale weight and influence).

The researcher submitted the final collection of 15 items (5 items per subscale) to a confirmatory factor analysis (CFA) to identify the initial factor structure and to estimate construct validity of the COSET. The CFA technique analyzes a priori measurement models in which both the number of factors and their correspondence with the indicators are explicitly specified (Kline, 2011). Data were analyzed with Amos 21.0, a software program for structural equation models and CFAs. Model identification was achieved by fixing the variance of each factor to 1.0. Item scores were only allowed to load on their intended latent factor, factors were allowed to correlate, but error terms were not. The goodness of fit of each of the models was based on the chi-square ($\chi^2$) statistic and several additional indices, including the minimum value of the discrepancy-C divided by the degree of freedom (CMIN/df), comparative fit index (CFI), the non-normed fit index (NNFI), and the root mean square error of approximation (RMSEA). The
researcher used 2.00 for the CMIN/df as the cutoff that higher values indicate an inadequate fit (Schumacker & Lomax, 2004). In addition, values less than .06 for the RMSEA (Hu & Bentler, 1999) and values above .95 for the CFI and NNFI (Hu & Bentler, 1999) generally indicate a good fit to the data. Loading of the items on the latent factor were estimated using a maximum-likelihood (ML) analysis. As mentioned previously, items were deleted if they produced scores that had factor loadings less than .40 (Mullan, Markland, & Ingledew, 1997) and were not needed for reasons of content sampling. Modification indices were also examined to locate potential cross-loading items. Items were considered for deletion if the results suggested an overall improved fit (Cheung & Rensvold, 2002).

Two models were compared in the confirmatory factor analysis: a one-factor model and three-factor model. Because the COSET is intended to produce a total test score that is a measure of overall supervisory quality, a one factor model was appropriate for consideration. Also, because the instrument assesses supervisory behaviors in three separate environments, a three factor model was considered an appropriate possibility. The first model tested was a one-factor model in which all the items were loaded to one single factor. In the second model, the items were evenly separated into three factors.

A subsequent reliability analysis was conducted to determine internal consistency of the final 15 COSET items and three subscales of the COSET. Additionally, a correlation analysis was conducted to examine the strength of
relationship between factors, with a goal of having moderate but not too strong inter-scale correlations.

Participants

The target population of this study was counselor educators and clinical supervisors across the United States. The sample was drawn from the target population. The researcher implemented a convenience sample technique for this study. To recruit participants, the researcher contacted counselor educators and clinical supervisors in two settings: university settings and mental health agencies. The researcher also distributed an invitation to counselor educators and counseling supervisors who are registered on CESNET (the listserv for counselor educators and counseling supervisors created by Kent State University). The invitation included information regarding the purpose of the study as well as an electronic link to the initial instrument in online version. Recipients who voluntarily participated in the study were instructed to go to a survey portal consisting of the consent form, instructions, demographics, and the initial instrument.

Initially, 115 participants responded to the initiation for the study; ultimately, the researcher deleted 15 cases with missing values and 7 cases with irregular response patterns. The final participants included 93 clinical supervisors between the ages of 26 and 74 years, with a mean of approximately 5 years of supervisory experience. Demographic characteristics of the sample are summarized in Table 2.1. Of the 93 subjects, 31 (33.3%) were male and 62 (66.7%) were female. Additionally, the sample included 77 (82.8%)
White/European/Caucasian Americans, 10 (10.8%) African or Black Americans,
2 (2.2%) Asian American or Pacific Islanders, and each of the following ethnic
groups had 1 (1.1%) supervising respondent: Hispanic or Latino Americans,
Native Americans and multiracial, and international.

Table 2.1

**COSET Gender by Race Sample Frequencies**

<table>
<thead>
<tr>
<th>Race</th>
<th>Number</th>
<th>Percent</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>African/Black American</td>
<td>10</td>
<td>10.8</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td>Asian American/Pacific Islander</td>
<td>2</td>
<td>2.2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>European/White/Caucasian American</td>
<td>77</td>
<td>82.8</td>
<td>27</td>
<td>50</td>
</tr>
<tr>
<td>Latino/Latina /Hispanic American</td>
<td>1</td>
<td>1.1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Native American</td>
<td>1</td>
<td>1.1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Multiracial</td>
<td>1</td>
<td>1.1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>International</td>
<td>1</td>
<td>1.1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>93</td>
<td>100.0</td>
<td>31</td>
<td>62</td>
</tr>
</tbody>
</table>

Table 2.2 outlines the professional roles indicated by the participants.

Fifty-eight of the participants identified themselves as Licensed Professional
Counselors (LPC), 19 as doctoral students in counselor education, 3 as Licensed
Marriage and Family Counselors (LMFC), 3 as counselor educators, 2 as
Residents in Counseling, 2 as Clinical Social Workers, 2 as Counseling or
Clinical Psychologists, 2 as doctoral students in counseling or clinical psychology,
1 as Licensed Professional Counselor Associate (LPCA), and 1 as Licensed Substance Abuse Treatment Practitioner (LSATP).

Table 2.2

*Distribution of COSET Sample Professional Roles*

<table>
<thead>
<tr>
<th>Professional Role</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clinical Social Worker</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Counseling/Clinical Psychologist</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Counselor Educator</td>
<td>3</td>
<td>3.2</td>
</tr>
<tr>
<td>Doctoral Student in Counselor Education</td>
<td>19</td>
<td>20.4</td>
</tr>
<tr>
<td>Doctoral Student in Counseling/Clinical Psychology</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Licensed Marriage and Family Counselor (LMFC)</td>
<td>3</td>
<td>3.2</td>
</tr>
<tr>
<td>Licensed Professional Counselor (LPC)</td>
<td>58</td>
<td>62.4</td>
</tr>
<tr>
<td>Resident in Counseling</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Other</td>
<td>2</td>
<td>2.2</td>
</tr>
<tr>
<td>Total</td>
<td>93</td>
<td>100.0</td>
</tr>
</tbody>
</table>

Table 2.3 outlines the participants’ years of supervision experience. Fifteen of the participants had less than one year of supervision experience, 41 had one to five years, 19 had six to 10 years, five had 11 to 15 years, three had 16 to 20 years, and 10 over 21 years (Table 3).
Table 2.3

*Distribution of COSET Sample Supervision Experience*

<table>
<thead>
<tr>
<th>Years</th>
<th>Number</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Less than 1 year</td>
<td>15</td>
<td>16.1</td>
</tr>
<tr>
<td>1 ~ 5 years</td>
<td>41</td>
<td>44.1</td>
</tr>
<tr>
<td>6 ~ 10 years</td>
<td>19</td>
<td>20.4</td>
</tr>
<tr>
<td>11 ~ 15 years</td>
<td>5</td>
<td>5.4</td>
</tr>
<tr>
<td>16 ~ 20 years</td>
<td>3</td>
<td>3.2</td>
</tr>
<tr>
<td>More than 21 years</td>
<td>10</td>
<td>10.8</td>
</tr>
<tr>
<td>Total</td>
<td>93</td>
<td>100.0</td>
</tr>
</tbody>
</table>

**Procedures**

The researcher obtained approval to conduct research on human subjects from the College of William and Mary Institutional Review Board (IRB). Participants were informed of the purpose and procedures involved in the study and their rights as participants (Appendix A). The researcher’s phone number and email were provided to participants to assist them if clarification or assistance was necessary. Participants were also informed that each questionnaire was coded using a number-coded system that ensured that participants’ identities would be protected. Electronic documents were stored in a secure website that requires a personal secure password to log in, and the password is only recognized by the researcher. Once participants agreed to participate in the study, the researcher asked them to complete the online COSET questionnaire including a survey of
demographic information and the initial instrument (Appendix B). The participants were allowed as much time as necessary to complete all items. The COSET administration time was estimated to be 15 minutes. The data collection process continued during a month period (March - April, 2013). The researcher collected data only through the online survey. The Qualtrics online survey program converted the survey responses to the SPSS database.
CHAPTER THREE

RESULTS

This chapter presents the results of the data analyses with the goal of developing and evaluating the psychometric properties of the Chae Optimal Supervision Environment Test (COSET). A review of the sample data will be reported. Then, the results of preliminary data analyses, the report of the findings of the factor analysis, the results of descriptive statistics, including basic descriptive statistics, internal consistency estimates, and correlation analysis among three components of the COSET will be presented. Finally, goodness of fit statistics for two models of factor structure of the COSET will be reported.

Sample Data

A total of 93 clinical supervisors participated in this study. The mean of the sample was approximately 44 years, 11 months and ranged from 26 to 74 years. Demographic characteristics of the sample are summarized in Table 1. Sixty-seven percent of the sample was female, and 17% indicated membership in an ethnic minority group. Table 2 outlines the professional role indicated by the sample. Sixty-two percent of the sample identified themselves as Licensed Professional Counselors (LPC), 20% as doctoral students in counselor education, and 18% as other professional roles. Forty-four percent of the sample reported one to five years of supervision experience, 20% reported six to 10 years of experience, 16% reported less than one year of experience, 11% reported more than 21 years of experience, and 9% reported 11 to 20 years of experience (Table 3).
Preliminary Data Analyses

In order to examine whether demographic variables (i.e., age, race, and gender) systematically influence the score of the COSET, a three-way analysis of variance (ANOVA), with age, gender, and ethnic group as independent variables was conducted. Because of the small group sizes, the ethnic minority categories were combined, resulting in two levels of race/ethnicity: Caucasian/European/White American ($n = 77$) and minority ($n = 16$). For the purpose of analysis, age was coded into five groups: 26 to 30 years ($n = 13$), 31 to 40 years ($n = 26$), 41 to 50 years ($n = 17$), 51 to 60 years ($n = 21$), and 61 and above ($n = 14$). Two participants did not indicate their age in the survey.

Analysis of variance showed no significant main effects at the $p < .05$ level for age: $[F (4, 74) = 2.34, p = .06]$, gender, $[F (1, 74) = 2.53, p = .146]$, or race/ethnicity, $[F (1, 74) = .41, p = .52]$. There were also no significant interactions. Because of these results, the sample was treated as one homogeneous group, regardless of age, gender, or race/ethnicity.

Validity

Internal Structure of the COSET

*Factor Analysis.* An exploratory factor analysis (EFA) was conducted using a principle component analysis (PCA) and a Varimax rotation to reduce the data set into a smaller number of variables and to reveal the underlying factor structure of the COSET. Squared multiple correlations were used as the initial communality estimates, and the communality estimates were iterated. The number of factors to be extracted was determined by eigenvalues of greater than
1.0, inspection of the scree plot (Cattell, 1966), and extraction criteria of .40 (Kline, 1986).

The 15 COSET items were subjected to an exploratory factor analysis (EFA). The significance of Bartlett’s test of sphericity, \( \chi^2 (105) = 751.76, p < .001 \), and the size of the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy, .87, showed that the 15 COSET items had an adequate common variance for factor analysis. The communalities were all above .3 (Table 3.1); further confirming that each item shared some common variance with other items.

Table 3.1

*Pattern Matrix and Communalities for the Three Factor Model of COSET*

<table>
<thead>
<tr>
<th>Item</th>
<th>EE</th>
<th>LE</th>
<th>PE</th>
<th>Communi ty</th>
</tr>
</thead>
<tbody>
<tr>
<td>My supervisee felt “safe” during our supervisory sessions.</td>
<td>.82</td>
<td>.70</td>
<td></td>
<td></td>
</tr>
<tr>
<td>My supervisee interacted with me in a genuine manner.</td>
<td>.82</td>
<td>.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Our supervisory relationship was characterized by a sense of mutual trust.</td>
<td>.81</td>
<td>.75</td>
<td></td>
<td></td>
</tr>
<tr>
<td>There was a positive atmosphere during our supervisory sessions.</td>
<td>.77</td>
<td>.64</td>
<td></td>
<td></td>
</tr>
<tr>
<td>My supervisee and I shared mutual respect as part of our supervisory relationship.</td>
<td>.81</td>
<td>.73</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I was aware of and sensitive to my supervisee's training needs.</td>
<td>.62</td>
<td>.57</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I matched my supervision approach to my supervisee's level of experience.</td>
<td>.85</td>
<td>.81</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I tailored supervision to my supervisee’s level of</td>
<td>.89</td>
<td>.82</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
competence.

I valued my supervisee's explanations about clients' behaviors.  

I modeled appropriate personal and professional boundaries.  

I acknowledged when my supervisee had made progress towards supervision goals.  

I consistently provided evaluation feedback to my supervisee.  

I was aware of and sensitive to the supervision evaluative process.  

I provided evaluative feedback based on observations of my supervisee’s performance.  

I regularly monitored my supervisee’s ethical behaviors.

<table>
<thead>
<tr>
<th></th>
<th>Factor Loadings</th>
<th>Communalities</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. EE = Emotional Environment; LE = Learning Environment; PE = Power Environment. 

Based on these criteria, three factors emerged with eigenvalues of greater than 1.0 after 5 iterations, accounting for 66.2% of the overall variance. The COSET items loaded onto three factors that correspond to Emotional Environment (EE), Learning Environment (LE), and Power Environment (PE). Each factor equally contained 5 items. Table 3.1 presents the factor loadings from the pattern matrix and the communalities for the three-factor model of the COSET. The first factor, the EE, accounted for 23.9% of the variance, with factor loadings for this factor ranging from .77 to .82. The second factor, the LE, accounted for 20.7% of the variance with the factor loading on this factor ranging
from .47 to .89. The last factor, the PE, accounted for 21.6% of the variance with factor loadings on this factor ranging from .64 to .80. The total variance was distributed approximately equally to the three factors of COSET.

**Descriptive Statistics**

Analysis of the data provided basic descriptive results of the supervisors’ scores on the COSET. Overall, participants scored a mean of 52.49 (SD = 5.39). The minimum and maximum possible overall COSET score is 15.00 and 60.00, respectively. The EE scores ranged from 11.00 to 20.00, with a mean of 17.69 (SD = 2.33). The LE scores ranged from 13.00 to 20.00, with a mean of 17.58 (SD = 2.18). The PE scores ranged from 13.00 to 20.00, with a mean of 17.23 (SD = 2.18). The minimum and maximum possible score for each COSET factor is 5.00 and 15.00, respectively. Means and standard deviations for each COSET item, as well as item-scale correlations, are shown in Table 3.2.

To determine the internal consistency of the COSET, Cronbach’s coefficient alpha was computed on the 15 items of the COSET total scale and each of the factors derived from the exploratory factor analysis. The overall total test alpha coefficient for this sample was .90. Alpha coefficients for the three factors were .89 (Emotional Environment), .86 (Learning Environment, and .84 (Power Environment). Wasserman and Bracken (2013) suggested that scales intended for research applications should minimally be reliable at a level of .70, and preferably .80. The reliability scores for this sample were considered excellent since the reliability scores for both the overall scale and each subscale...
were well above the preferred .80 for scale reliability. Overall, the reliability analyses provide support for the COSET as a reliable instrument.

Table 3.3 presents the intercorrelations for the COSET subscales and the total scale scores. As can be seen in Table 3.3, the three COSET subscales correlated to a moderate to strong degree with the COSET total scale. Subscale to total scale intercorrelations coefficients ranged from a low .77 for EE and total scale, to a high of .85 for LE and total scale. The EE subscale correlated with the LE subscale at .47 \( (p < .01) \), which means the two scales have 22% of their variance common. Given the EE subscale reliability of .89 and variance held in common with LE subscale, 67% of its variance is unique and separate from the LE subscale. Also, the LE subscale reliability is .86; thus, 64% of its variance is unique and separate from the EE subscale. The EE subscale correlated with the PE subscale at .37 \( (p < .01) \), which means the two scales share 17% of their variance. Given the EE subscale reliability of .89 and variance held in common with PE subscale, 72% of its variance is unique and separate from the PE subscale. Also, the PE subscale reliability is .84; thus, 67% of its variance is unique and separate from the EE subscale. The LE subscale correlated with the PE subscale at .60 \( (p < .01) \), which means the two scales have 36% of their variance common. Given the LE subscale reliability of .86 and variance held in common with PE subscale, 50% of its variance is unique and separate from the PE subscale. Also, the PE subscale reliability is .84; thus, 48% of its variance is unique and separate from the LE subscale. Overall, these findings suggest that the three indices of the COSET are related but not sufficiently explained by one score alone. The results
support the discriminant validity of the factor scores and suggest that interpreting the total test and the three subscales is acceptable.

Table 3.2

*Means, Standard Deviation, and Item-Scale Correlations of COSET*

<table>
<thead>
<tr>
<th>Item</th>
<th>M</th>
<th>SD</th>
<th>Scale Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>My supervisee felt “safe” during our supervisory sessions.</td>
<td>3.48</td>
<td>.54</td>
<td>.58</td>
</tr>
<tr>
<td>My supervisee interacted with me in a genuine manner.</td>
<td>3.55</td>
<td>.58</td>
<td>.66</td>
</tr>
<tr>
<td>Our supervisory relationship was characterized by a sense of mutual trust.</td>
<td>3.48</td>
<td>.60</td>
<td>.70</td>
</tr>
<tr>
<td>There was a positive atmosphere during our supervisory sessions.</td>
<td>3.58</td>
<td>.52</td>
<td>.60</td>
</tr>
<tr>
<td>My supervisee and I shared mutual respect as part of our supervisory relationship.</td>
<td>3.59</td>
<td>.54</td>
<td>.68</td>
</tr>
<tr>
<td>I was aware of and sensitive to my supervisee's training needs.</td>
<td>3.49</td>
<td>.50</td>
<td>.67</td>
</tr>
<tr>
<td>I matched my supervision approach to my supervisee's level of experience.</td>
<td>3.46</td>
<td>.60</td>
<td>.73</td>
</tr>
<tr>
<td>I tailored supervision to my supervisee’s level of competence.</td>
<td>3.45</td>
<td>.60</td>
<td>.64</td>
</tr>
<tr>
<td>I valued my supervisee's explanations about clients' behaviors.</td>
<td>3.58</td>
<td>.52</td>
<td>.68</td>
</tr>
<tr>
<td>I modeled appropriate personal and professional boundaries.</td>
<td>3.59</td>
<td>.52</td>
<td>.66</td>
</tr>
<tr>
<td>I acknowledged when my supervisee had made progress towards supervision goals.</td>
<td>3.55</td>
<td>.52</td>
<td>.60</td>
</tr>
<tr>
<td>I consistently provided evaluation feedback to my supervisee.</td>
<td>3.31</td>
<td>.53</td>
<td>.64</td>
</tr>
</tbody>
</table>
supervisee.

I was aware of and sensitive to the supervision evaluative process. 

I provided evaluative feedback based on observations of my supervisee’s performance. 

I regularly monitored my supervisee’s ethical behaviors. 

3.47 .54 .63 

3.43 .60 .58 

3.46 .58 .71 

**Note.** M = mean; SD = standard deviation

**Table 3.3**

*Means and Intercorrelations for EE, LE, and PE*

<table>
<thead>
<tr>
<th>Variables</th>
<th>M</th>
<th>SD</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Emotional Environment</td>
<td>17.69</td>
<td>2.33</td>
<td></td>
<td>0.47*</td>
<td>0.37*</td>
<td>0.77*</td>
</tr>
<tr>
<td>2. Learning Environment</td>
<td>17.58</td>
<td>2.18</td>
<td></td>
<td>0.60*</td>
<td></td>
<td>0.85*</td>
</tr>
<tr>
<td>3. Power Environment</td>
<td>17.23</td>
<td>2.18</td>
<td></td>
<td></td>
<td>0.81*</td>
<td></td>
</tr>
<tr>
<td>4. Total Scale</td>
<td>52.49</td>
<td>5.39</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. M = mean; SD = standard deviation*

* p < .01

**Measurement Model**

A confirmatory factor analysis was used to compare the estimate of fit for each of two measurement models: a one factor model and a three factor model.

For the one factor model, there was one latent variable, the *Optimal Supervision Environment*, which had 15 indicators. For the three factor measurement model, three latent variables, Emotional Environment (EE), Learning Environment (LE), and Power Environment (PE), each had 5 indicators. The three latent variables were allowed to correlate as shown in the correlation analyses among the factors.
The adequacy of measurement and structural model fit was based on the chi-square ($\chi^2$) statistic and several additional indices, including the minimum value of the discrepancy-C divided by the degree of freedom (CMIN/df), comparative fit index (CFI), the non-normed fit index (NNFI), and the root mean square error of approximation (RMSEA). Various cutoffs ranging from 2 to 5 have been suggested for CMIN/df (e.g., Byrne, 1989; Carmines & McIver, 1981; Marsh & Hocevar, 1985; Schumacker & Lomax, 2004). In this study, the researcher used 2.00 as the cutoff, with higher values indicating an inadequate fit (Schumacker & Lomax, 2004). In addition, values less than .06 for the RMSEA (Hu & Bentler, 1999) and values above .95 for the CFI and NNFI (Hu & Bentler, 1999) indicated a generally good fit to the data.

The models were tested using maximum likelihood estimation. Diagrams of the one factor model and the three factor model showing the standardized estimations of the paths are presented in Figure 3.1 and 3.2, respectively. Standardized regression weights on each item were > .40, highlighting good factor loading. The hypothesized one-factor model of COSET was examined and the data showed a poor fit to the model according to the approximate fit indices: $\chi^2 (90, N = 93) = 341.473, p < .001; \text{CMIN/df} = 3.79; \text{CFI} = .64; \text{NNFI} = .58; \text{and RMSEA} = .17$. On the other hand, the model fit statistics for the three factor model of COSET indicated a very good fit to the data, $\chi^2 (87, N = 93) = 116.33, p = .02; \text{CMIN/df} = 1.34; \text{CFI} = .96; \text{NNFI} = .95; \text{and RMSEA} = .06$. Item scores loaded strongly on the intended factor. Modification indices were inspected, and no items appeared to cross-load. Therefore, the results of fit indices for the two
models suggest that the three factor model is superior to the one-factor model.

The confirmatory factor analysis also supports the interpretation of the three respective scales, as well as the total test score as an overall measure.

Figure 3.1

*Three-Factor Model of COSET*
Figure 3.2

One-Factor Model of COSET
Conclusion

The goal of this study was to develop and evaluate the psychometric properties of the Chae Optimal Supervision Environment Test (COSET). Results based on 93 counselor educators and clinical supervisors indicated that the instrument yields three factors: Emotional Environment, Learning Environment, and Power Environment. Sixty-six percent of the variable was explained. The COSET demonstrated high internal consistency with an overall Cronbach’s alpha of .90. The three-factor model met all the model fit statistics criteria. The results demonstrate that the COSET has reliable psychometric properties for use in supervision research and clinical settings.
Chapter 4

DISCUSSION

The primary purpose of this study was the construction and initial validation of the Chae Optimal Supervision Environment Test (COSET). The five phases of scale development provided preliminary evidence of reliability and validity for the COSET. The results are largely supportive of the COSET as a scale to assess supervisors’ creation of optimal supervision environments.

Rationale

Bernard and Goodyear (2009) emphasized the importance of clinical supervision in the training and development of professional counselors; however, numerous studies have reported that supervisees feel anxious, traumatized, exploited, and doubtful of their abilities as counselors in response to negative supervisory experiences (e.g., Ellis, 2001; Gray et al., 2001; Greer, 2002; Jernigan et al., 2010; Ladany et al., 1999; Magnuson et al., 2000; Watkins, 1997). In order to prevent these negative supervision experiences, considerable research interest has been focused on the effects of the supervision environment on supervisees (Bernard & Goodyear, 2009). Unfortunately, the lack of sufficiently valid and reliable supervision-specific instruments has hindered the quality in supervision research. In response to the perceived need for quality instrumentation, this researcher conducted a careful review and evaluation of existing supervision models and instruments for content sampling and psychometric soundness, and concluded that extant scales were inadequate to address the components of the optimal supervision environment in a satisfactory manner. Given the recognition
that existing measures have many limitations, the purpose of this study was to develop a psychometrically and theoretically sound supervision instrument to assess the quality of the optimal supervision environment from the supervisor’s perspective.

**Scale Development of the COSET**

The construction and field-testing of the COSET took place in five phases. Drawing from the current supervision literature and applying two developmental theories, constructive developmental theory (Kegan, 1982, 1994) and adult attachment theory (Bowlby, 1998), the researcher identified three core elements that comprise the optimal supervision environment. These three core elements (i.e., the Emotional Environment, the Learning Environment, and the Power Environment) served as the basis for creating a comprehensive test development blueprint. Approximately 300 items were generated through multiple iterations of item writing, editing, review, and modification. A pilot study was conducted to improve the clarity and quality of items as well as to reduce their overall number, which was then followed by an expert panel review of the COSET items. Lastly, the researcher conducted a field study and collected data from 93 counseling supervisors.

**Descriptions of the COSET**

The COSET assesses supervisors’ perceived ability to create an optimal supervision environment through three subscales: (a) the Emotional Environment (EE), (b) Learning Environment (LE), and (c) Power Environment (PE). The COSET is designed to be administered by counselor educators, supervisors, and
supervisors in-training in group or individual test administration venues.

Although, the COSET does not have administration time limits, the instrument can be administered in approximately five minutes. The COSET contains 15 Likert-type self-report items with four response options, and no neutral option. Each subscale contains five items (see Appendix C). Administration of the COSET results in four scores: (a) the total COSET score; (b) the EE score; (c) the LE score; and (d) the PE score. The raw scores of three subscales are combined to create the overall raw COSET score. The score for the total COSET ranges from 15 to 60; the three subscale scores range from 5 to 20.

**Emotional Environment Subscale**

The content of the items on the EE subscale describe the supervisor’s understanding of supervisees’ emotional needs and ability to create a healthy supervisory relationship that promotes counselor development. Items on this subscale captured the notion that supervisors should initiate the supervisory relationship by appreciating the emotional needs of supervisees and creating an environment that allows supervisees to feel safe and supported.

Watkins (2010) suggested that supervisors need to establish the supervisory relationship as a holding environment (Winnicott, 1965) by creating a safe environment for the supervisee in which trust, consistency, and dependability permeate every facet of the supervisory relationship. Researchers have applied attachment theory to the supervisory relationship and proposed that supervisors should function as a secure, safe base from which supervisees can explore and develop their counseling skills and professional identity (Palomo et al., 2010;
Pistole & Watkins, 1995; Watkins, 1997). Along the same line, White and Queener (2003) proposed that supervisors’ ability to form close relationships and to feel intimate in relationships are predictive of a strong supervisory working alliance. Riggs and Bretz (2006) also suggested that the supervisees’ perception of supervisors’ attachment style has a positive influence on the supervisory relationship. The EE subscale items support this theoretical dimension by assessing the supervisors’ perceived capacity to provide a safe and secure supervisory relationship.

Many studies have discussed the importance of providing a safe and supportive space for supervisees to discuss cultural issues in supervision. For example, Dressel, Consoli, Kim, and Atkinson (2007) used the Delphi method to have 21 supervisors rank elements of successful and unsuccessful multicultural supervision. The raters identified several supervisory behaviors associated with successful multicultural supervision such as: creating a safe environment (i.e., nonjudgmental and supportive) for discussion of multicultural issues, listening to and demonstrating genuine respect for supervisees’ ideas about how culture influences the clinical interaction, providing openness and genuineness, and communicating acceptance of and respect for supervisees’ culture and perspective.

The content of the EE subscale also reflects supervisors’ perceived ability to build trust and mutual respect in supervisory relationships. The items contributing to the EE subscale include: (a) My supervisee felt “safe” during our supervisory sessions; (b) My supervisee interacted with me in a genuine manner; (c) Our supervisory relationship was characterized by a sense of mutual trust; (d)
There was a positive atmosphere during our supervisory sessions; and (e) My supervisee and I shared mutual respect as part of our supervisory relationship. These five items reflect the core emotional elements in the supervisory relationship related to a safe, positive, and supportive environment that allows supervisees to disclose their professional and personal struggles in counseling.

**Learning Environment Subscale**

The items of the LE subscale assess supervisors’ perceived ability to understand supervisees’ learning needs and to intervene during supervision according to the supervisees’ developmental level. Competent supervisors are skilled educators who impart their counseling knowledge and skills by matching supervision interventions according to their supervisees’ cognitive developmental levels (Borders, 1989).

Research studies demonstrate the importance of creating a learning environment that supports and challenges the supervisees’ cognitive developmental level (e.g., Borders, 1989; Borders & Fong, 1989; Borders et al., 1986; Ladany et al., 2001; Lovell, 1999). Barrett and Barber (2005) suggested that supervisors need to match the supervisees’ developmental level to supervisory interventions according to the supervisees’ developmental needs. They contend that appropriately matched supervisory interventions have the potential to limit negative interactions between supervisors and supervisees that result from incorrect assumptions about the needs of supervisees or inaccurate expectations of their behaviors. Magnuson et al. (2000) also stressed the importance of supervisors’ sensitivity to the supervisees’ developmental level.
Expecting supervisees to think and behave in ways that are beyond their developmental capacities may result in frustration and dissatisfaction. The items in the Learning Environment (LE) subscale measure the extent to which supervisors believe they understand supervisee's ways of learning and their ability to create a learning environment that promotes counselor development.

The items contributing to the LE subscale include: (a) I was aware of and sensitive to my supervisee's training needs; (b) I matched my supervision approach to my supervisee's level of experience; (c) I tailored supervision to my supervisee's level of competence; (d) I valued my supervisee's explanations about clients' behaviors; and (e) I modeled appropriate personal and professional boundaries. These items capture supervisors' awareness of supervisees' competence and experience level and then matching their supervision accordingly. The LE subscale assesses supervisors' perceived provision of a learning environment that matches their supervisees' way of learning.

**Power Environment Subscale**

The PE assesses the supervisor's perceived ability to understand the hierarchical, evaluative nature of the supervisory relationship and to create an evaluative environment that promotes counselor development. Evaluation and feedback are important roles for supervisors when monitoring the quality of professional services supervisees offer to clients. Such evaluation and feedback positions the supervisor as a gatekeeper for the profession, monitoring and facilitating supervisee growth and development (Bernard & Goodyear, 2009).
modeling effective feedback for supervisees (Freeman, 1985), and encouraging supervisees' self-evaluation (Farnill, Gordon, & Sansom, 1997).

Research has shown the importance of evaluation and constructive feedback in clinical supervision. For example, one study reported that supervisor's observation and feedback was considered as the most effective factor contributing to their skill development (Smith, 1984). In addition, Stoltenberg et al. (1998) emphasized the importance of matching feedback and evaluation to the developmental level of the supervisee. However, supervisors' inadequate provision of feedback and evaluation is the focus of most ethical complaints involved in the supervision relationship (Ladany et al., 1999).

The items contributing to the PE subscale include: (a) I acknowledged when my supervisee had made progress towards supervision goals; (b) I consistently provided evaluation feedback to my supervisee; (c) I was aware of and sensitive to the supervision evaluative process; (d) I provided evaluative feedback based on observations of my supervisee's performance; and (e) I regularly monitored my supervisee's ethical behaviors. These items address the importance of supervisors' provision of effective feedback and sensitive evaluation during supervision. The PE subscale provides an estimate of the evaluation process and its implications for supervisees.

**Psychometric Characteristics of the COSET**

**Reliability**

Reliability represents the percentage of variance in test scores that is a result of reliable variation, as opposed to variation that results from error. The
total scale and each COSET subscale were shown to have strong internal consistency. The COSET subscale scores had sufficient reliability for research purposes with alpha coefficients above .80 and .90 for total scales as recommended by Wasserman and Bracken (2013). The measures of internal consistency are inversely related to measurement error and therefore suggest that measurement error is minimized in the COSET, ranging from 11% to 16% for the subscales, and 10% for the total scale. In other words, approximately 90% of the total COSET variance is reliable, and only 10% is due to error. The estimates of internal consistency of the total COSET and its subscales suggest that examiners can expect examinee item responses to be consistent within scales. The high alpha coefficients also suggest that the COSET will likely perform reliably in future research and training applications.

Validity

The researcher used a principle component analysis (PCA) as the exploratory factor analysis (EFA) method to examine the initial factor structure and construct validity of this scale. Based on the results of the EFA, a three-factor solution appeared to describe the dimensions of optimal supervision environment. The final COSET contained 15 items with three subscales: the Emotional Environment (EE), the Learning Environment (LE), and the Power Environment (PE). Primary factor loadings for each of the three factors were clearly identified, and only one of the 15 items on the COSET had primary loading of less than .50 on its respective scale. The results of the EFA provide initial evidence in support of the COSET construct validity, in that it demonstrates
an interpretable underlying factor structure that coincides with the instrument’s theoretically-based blueprint.

In the present study, the final factor analysis performed on the 15-item COSET had a ratio of participants to items greater than 6:1, with items per factor and the majority of factor loadings greater than .60. Only two of the 15 items’ communalities were less than .60 (see Table 3.1). Worthington and Whittaker (2006) suggested that smaller samples may be adequate for the purposes of factor analysis if the analyses yield communalities of .60 or greater or there are at least four items per factor, and the factor loadings are greater than .60.

With regard to the total sample size for EFA, Gorsuch (1983) also recommended at least a 5:1 ratio of participants to items. This sample, therefore, satisfied Gorsuch’s recommended ratio and satisfied Worthington and Whittaker’s recommendation of items per factor and factor loading magnitudes. Additionally, the size of the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy was good (> .60) which further supports the appropriateness of the sample for this study.

Ellis and Ladany (1997) recommended the use of confirmatory factor analysis (CFA) in instrument development and the testing of an a priori factor model in supervision research. This study satisfied their recommendations by using the CFA to test the hypothesized three-factor structure of the COSET. The results indicated that the model’s goodness-of-fit with the data was good but not excellent; however, it still, satisfied all recommended criterion. One possible explanation for a less than excellent fit is that the field study employed a
relatively small sample. Because the exact sample size needed to perform a
reliable CFA is not well established (Kline, 2005), it is difficult to identify the
extent to which sample size affected the overall statistical fit of the model.

The CFA also provided support for the multidimensionality of the
supervisory construct. The three-factor model had slightly better fit indices than
the alternate proposed one-factor model. These results indicate that not only does
the COSET assess the nature of the optimal supervision environment, but also
demonstrates that the optimal supervision environment can be viewed as a
multidimensional construct.

The intercorrelations of the COSET subscales were moderate, suggesting
that the three factors of the COSET are related. This was to be expected, since
they are each part of counseling supervision but not sufficiently explained by a
single total test score. The results support the discriminant validity of the three
factor scores; importantly, the minor differences between the one and three factor
CFA solutions suggest that interpreting either or both the total test and the three
subscales would be appropriate. Since the COSET was based on a comprehensive
blueprint that guided instrument development, it seems likely that the three-factor
structure for the COSET will also be supported with future samples of clinical
supervisors.

A three-way analysis of variance (ANOVA), using supervisees’ age,
gender, and ethnic group as independent variables and the COSET score as a
dependent variable, was conducted and found no significant effects for all
independent variables. This finding demonstrates that the scale performs
consistently across demographic groups, thus suggesting that the COSET scores are affected more by supervisors' ability rather than the characteristics of the participants.

The researcher's goal for developing the COSET was to address the critical need for an empirically supported instrument for supervisors to use in the assessment of an optimal supervision environment. Because supervisors are an essential element in the development and training of competent counselors, it is important for supervisors to understand supervisory competence. By using the COSET, it is proposed that supervisors can explore and evaluate the extent to which they create an optimal supervision environment. Results from the COSET can also be used to improve supervision by identifying supervision environments that are problematic. The findings of the current study provide an important first step toward validation, but further efforts to assess the psychometric properties of the COSET are needed.

Implications to Counselor Education and Supervision

The psychometric support for the COSET has important implications for supervisors, counselor educators, and supervisors in-training. The COSET is designed to be used as a self-report assessment of a supervisor's perceived ability to create an optimal supervision environment. Using the self-report format, counselor educators and supervisors may consider using the COSET as a tool for self-assessment and for program evaluation. Because the scale is brief and easily administered, the COSET can be given to supervisors as an efficient way of self-evaluation. They can use the COSET subscales to determine supervisory
functions that work effectively and the functions that may need additional attention. The COSET also may be used as a measure for supervisors’ developmental growth as a professional.

Counselor educators can also use the COSET to monitor and provide feedback to supervisors on their supervisory behaviors. This scale might be useful as a tool for measuring novice supervisors’ understanding of their supervisory abilities. Comparing the COSET scores with other objective ratings of performance may offer valuable information regarding the ability of novice supervisors to engage in accurate self-reflections. Thus, the COSET allows counselor educators and supervisors to help the supervisors-in-training more accurately understand their professional development.

The COSET looks to be an ideal instrument for use in supervision research due to its psychometric properties and the concise and efficient test administration process. The total COSET score and its subscales can be utilized as a predictor or dependent variable. As a predictor variable, the COSET may have significant utility in researching counseling or supervision outcomes that are predicted by supervisor scores or patterns of scores on the instrument. For example, supervisors who demonstrate high COSET scores might be predicted to also demonstrate higher supervisee satisfaction scores, higher supervisee developmental level, or lower negative supervision experience. As a dependent variable, the COSET may have significant utility in supervision evaluation or supervision training program evaluation. The COSET may be useful for evaluating the extent to which a supervisor’s competency improved as a result of
supervision training. Also, the COSET would be valuable for tracking the developmental nature of the supervisor development during supervision training or counselor education program. For example, the COSET can be administered at the beginning of the supervision training program as a baseline, and for the duration of the program, the COSET can be utilized to monitor the supervisees’ development as a supervisor.

Limitations

Despite the numerous strengths of this study, there are also some limitations. Although the results reported in this study provide consistent evidence of reliability and construct validity to support the adequacy of the COSET, it is essential that the scale be used within the context of the study’s limitations. First, because the results were based on a convenience sample, results must be interpreted with caution. The study did not include any geographic or socioeconomic data in the survey, so the generalizability of the supervisor data may be somewhat limited. Demographics revealed the most common participant to be a white (83%), female (67%) supervisor identified as a Licensed Professional Counselor (62%), which mirrors those of clinical supervisors, especially ethnicity and gender. However, there is no body demographic data available that accurately reflects the “norm” of the population of clinical supervisors to which the researcher could compare the demographics of sample in this study. It should be noted that the sample of 93 participants includes wide age (26 to 74 years) and supervision experience (1 to 21 years and more) ranges, which implies that the sample has some large range of demographic data. Future
research efforts with the COSET should include additional questions on demographics regarding geographic and socioeconomic information, as well as a more diverse sample.

Additional limitations rooted in the sampling method may restrict the representation and generalization of the findings. Because the sampling method employed in this study did not permit tracking of invitees who responded and those who declined to respond, it is impossible to determine if there was significant difference between the two groups. For the purpose of this study, the researcher used the CESNET to distribute invitations to counselor educators and counseling supervisors and also made personal contacts for participation and/or distribution of the field study survey. Therefore, only participants registered on the CESNET or those invited through personal contact participated in the study, and this may have significantly limited the representation and generalization of the findings.

The second limitation of the current study is sample size. After accounting for missing data, 93 participants were included in the data analyses. There is a general agreement that larger sample sizes result in more stable correlations among variables and, therefore, greater potential for replication in validity estimates (Worthington & Whittaker, 2006). Riese, Waller, and Comrey (2000) conducted a literature review on scale development, and concluded that most general rules of thumb regarding to minimum sample size are not useful. They stated that when communalities are high and factors are well defined, sample sizes of 100 are often adequate; but when communalities are low, the number of
factors is large, and the number of indicators per factor is small, even a sample size of 500 may be inadequate. As mentioned earlier, Worthington and Whittaker (2006) provided a more specific guideline, suggesting that smaller samples may be adequate if factor yields communalities of .60 or greater or at least 4:1 items per factor and factor loading greater than .60. With regard to the total sample size for EFA, Gorsuch (1983) recommended at least a 5:1 ratio of participants to items. Although, the final factor analysis performed on the 15-item COSET had a ratio of more than 6:1 of participants to items, with strong communalities and factor loadings, the sample was slightly short of Riese et al. suggested size of 100. The COSET would benefit from additional studies with larger samples.

As a third limitation, the COSET, like other self-reported instruments, faces an issue of social desirability effect (Gall et al., 2007). Social desirability can create difficulties in research, particularly in psychological research, because results can often be affected by participants’ desire to be seen as acceptable by the researchers. To address concerns about social desirability, some measurements include a social desirability scale to the questionnaire. However, for the purpose of this study, a social desirability scale was not used. The researcher had placed a stronger emphasis on the scale length and parsimony than social desirability. Also, scale items were carefully developed not to make participants feel defensive; thus, it appeared that the risk of participants’ responses being influenced by social desirability bias was particularly low. In a future study, the COSET could benefit from a supplementary COSET scale that measures the same construct (i.e., supervisors’ ability to create an optimal supervision environment) but from the
supervisees' perspective. This pairing of content would allow researchers to better determine if social desirability had an effect on the COSET scores by comparing the perceptions of supervisors and supervisees across the findings of both instruments.

Lastly, there are some limitations in the scale development and investigation of technical adequacy. For example, the researcher did not investigate the test-retest reliability (i.e., stability) or criterion-related validity of the COSET. Bracken (1993) underlined the importance of demonstrating both internal consistency and stability for tests that are used for placement decisions in educational or psychological settings. Wasserman and Bracken (2013) stated that test-retest reliability and internal consistency are influenced by different types of error, making it is possible for one to be different from the other. Because supervision is an interpersonal relationship, one must question how stable the COSET would be expected to be across time and supervisees. Additionally, according to Wasserman and Bracken: “The accumulation of external evidence of test validity becomes most important when test results are generalized across contexts, situations, and populations and when the consequences of testing reach beyond the test’s original intent” (p. 66). Additional validation efforts using contrasted groups of supervisors who are known to be “good” or “poor” supervisors would be useful for examining the sensitivity of the instrument, as well as the veracity of respondents’ self-reported perceptions. Finally, to better understand the construct of supervisory environments, the COSET needs to be
included with other assessments of supervisory efficacy to better define its
discriminant and convergent relationships.

As a newly developed instrument, the COSET warrants further
investigation of its technical adequacy and possible contribution to the field. By
conducting further reliability and validity studies, researchers will better
understand the instrument and its practical use in research and supervision
evaluation.

Directions for Future Research

The COSET would benefit from additional replicated and cross-validation
research with a larger sample. The sample size of this study was smaller than
ideal. Also, adding more demographic questions might increase the likelihood to
enhance external validity.

It is recommended that the construct validity of the COSET scale be
examined further. The validity of COSET would benefit from future studies with
participants responding to other supervision instruments (e.g., Ladany et al., 1996;
Lanning, 1986; Oik & Friedlander, 1992). For example, the Supervisory
Relationship Questionnaire (SQR; Palomo et al., 2010) measures the quality of
the supervisory relationship which is similar to the EE subscale of the current
instrument. Also, the Evaluation Process Within Supervision Inventory (EPSI;
Lehrman-Waterman & Ladany, 2001), a measure that examines evaluation
practices in clinical supervision, might be appropriate for use in confirming the
COSET’s concurrent validity. The COSET could also be administered with
measures of satisfaction in supervision to assess convergent validity. Scores
across the two constructs are anticipated to be significantly and positively correlated.

It would be ideal to develop a supervisee’s version of the COSET that measures the same construct but from the supervisees’ perspective. This alternative version could offer test users a contrasted view and, possibly, a more accurate understanding of the supervisors’ ability to create an optimal supervision environment. It might be practical for research purposes to develop the supervisee version using the same blueprint as the COSET to coincide with the supervisor version and having the same content, number of items, and subscales.

Conclusion

The COSET appears to be a highly reliable measure with preliminary support for its validity as a measure supervisor’s ability to create an optimal supervision environment. The theoretically derived COSET provides a new and unique measure in the supervision field where there has been a lack of empirical research. The initial validation of the COSET has demonstrated a three-factor model that identifies three important environmental domains of supervision. The three subscales include: the Emotional Environment, the Learning Environment, and the Power environment. The COSET demonstrates reliability well above the acceptable range for its intended informal applications, and initial estimates of validity suggest the scale has good preliminary validity. The validation efforts demonstrate that the COSET may be a useful tool in measuring supervisor's professional development as a researcher. Ongoing research in the area of the
importance of supervision environments is encouraged in order to develop further understanding of this important aspect of counselor training and development.
Appendix A

INFORMED CONSENT

Dear Participant,

This is a dissertation research study investigating the supervisory environment. The research is conducted by Ki Chae, M.A., Dr. Charles “Rip” McAdams, and Dr. Bruce Bracken of the School of Education at the College of William & Mary. The study involves completing a brief survey as a clinical supervisor, asking about your perceptions of a recent supervisory experience. The survey will take about 10-15 minutes.

The purpose of this research is to better understand the quality of the supervisory environment and develop a reliable instrument that measures the supervisory environment from the supervisor's perspective. You will be asked about your thoughts and feelings towards a recent supervisee and the supervisory experience. The information you provide will be invaluable in helping us to better understand supervisory environment.

There is no expectation of discomfort expected from this survey research. The risks of participating are not greater than those ordinarily encountered in daily life; however, if you have negative feelings after completing these questionnaires and feel that you may need to talk with someone, please contact your campus based counseling center.

No identifying information is included in the questionnaires. Your responses are anonymous, and your individual responses cannot be identified. Only the researchers will see your responses. Your participation in the research is completely voluntary, and refusal to participate will involve no penalty or loss to you. You may terminate your participation at any time.

If you have additional questions or concerns about your rights as a participant, or are dissatisfied at any time with any aspect of this study, you may contact, anonymously if you wish, the two chairpersons of the W&M committees which supervise the treatment of study participants: Dr. Charles “Rip” McAdams at 757-221-2338 or crmcad@wm.edu, and Dr. Bruce Bracken at 757-221-1712 or babrac@wm.edu.

If you had read and agree to this informed consent, please select agree below. Thank you.
**Online COST Questionnaire**

**Appendix B**
What is the gender of this recent supervisee?
A supervisee of the opposite sex
A supervisee of the same sex

What is the professional identity of this supervisee?
Master's student in Counseling
Master's student in Counseling/Clinical Psychology
Master's student in Social Work
Master's student in School Counseling
PhD student in Counselor Education
PhD student in Counseling/Clinical Psychology
Unlicensed counselor with a Master's degree
Others (Please indicate below)

Approximately, how many years of counseling experiences does this supervisee have?
less than 1 year
1-3 years
3-5 years
5-10 years
10+ years
The following statements describe some of the ways a supervisor may feel about a supervisee or a supervisory experience. With the same supervisee in mind, to what extent do you agree or disagree with each of the following statements? Please click the column which matches your opinion most closely.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

I provided emotional support and encouragement during our supervisory sessions.

My behaviors and communications were authentic during our supervisory relationship.

My supervisee talked openly with me about sensitive counseling issues.

My supervisee felt “safe” during our supervisory sessions.

My supervisee interacted with me in a genuine manner.

My supervisee discussed personal struggles during our supervisory sessions.

My supervisee and I truly connected during our supervisory sessions.

Our supervisory relationship was characterized by a sense of mutual trust.

My supervisee expressed personal feelings during our supervisory sessions.

There was a positive atmosphere during our supervisory sessions.

I created a “safe place” for my supervisee to express feelings of being overwhelmed.

My supervisee and I shared mutual respect as part of our supervisory relationship.

My supervisee turned to me for comfort and reassurance.
The following statements describe some of the ways a supervisor may feel about a supervisee or a supervisory experience.

With this in mind, to what extent do you agree or disagree with each of the following statements?

Please click the column which matches your opinion most closely.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

1. I was appropriately assertive with my supervisee.
2. I was willing to stay late if my supervisee was dealing with a crisis.
3. My supervisee was willing to turn to me during times of need.
4. I was honest during our supervisory experience.
5. I felt comfortable expressing whatever was on my mind during supervision.
6. I shared my true impressions of my supervisee with him or her.
7. I challenged my supervisee to talk about feelings related to our supervision sessions.
8. I challenged my supervisee to identify personal biases.
9. I was available to my supervisee when emotional support was needed.
10. I adjusted my work schedule to be available to my supervisee.
11. I made sure that my supervisee was supported during crisis situations.
12. I was comfortable challenging my supervisee by creating tension.
13. I made an effort to make as much sense for my supervisee as was needed.
The following statements describe some of the ways a supervisor may feel about a supervisee or a supervisory experience.

With the same supervisee in mind, to what extent do you agree or disagree with each of the following statements?

Please click the column which matches your opinion most closely.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

I regularly followed up on cases we had discussed previously during our supervision.

I modeled problem-solving approaches.

I used supervision interventions that appropriately challenged my supervisee's developmental level.

I was a knowledgeable resource to my supervisee.

I suggested trainings or workshops my supervisee should attend.

I taught new counseling skills by demonstration.

I helped my supervisee see how interventions can affect clients.

I modeled appropriate personal and professional boundaries.

I provided opportunities for my supervisee to try new counseling techniques.

I was available when my supervisee had questions about clinical practice.

I provided regular feedback on professional development to my supervisee.

I provided challenges according to my supervisee's level of clinical competence.

I modeled ethical behaviors with my supervisee.
The following statements describe some of the ways a supervisor may feel about a supervisee or a supervisory experience. With the same supervisee in mind, to what extent do you agree or disagree with each of the following statements? Please click the column which matches your opinion most closely.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
</table>

I was sensitive to the power differential in our supervisory relationship.
I actively involved my supervisee in the performance evaluation process.
I welcomed comments about my supervision style.
I admitted mistakes I made in supervision.
I breached the power differential within our relationship with my supervisee.
I acknowledged when my supervisee had made progress towards supervision goals.
I empowered my supervisee to make independent case decisions.
I was open to receiving feedback on my own supervision performance.
My supervisee was comfortable asking about the supervision evaluation process.
My supervisee was comfortable talking about feelings of inadequacy as a counselor.
My supervisee was comfortable openly disagreeing with me.
I created supervision goals that matched my supervisee's developmental level.
I expected my supervisee to be prepared for supervision sessions.
The following statements describe some of the ways a supervisor may feel about a supervisee or a supervisory experience. With the same supervisee in mind, to what extent do you agree or disagree with each of the following statements? Please click the column which matches your opinion most closely.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>When needed, I confronted my supervisee on work ethics or performance.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I consistently provided evaluation feedback to my supervisee.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I addressed performance issues on an ongoing basis, not just in written evaluations.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I regularly balanced positive and negative evaluative comments.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I was straightforward with my supervisee when talking about his or her performance.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When providing evaluative feedback, I offered alternative ways to improve performance.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>When I thought my supervisee was wrong, I would let him or her know.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I was aware of and sensitive to the supervisee's evaluative process.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I provided evaluative feedback based on observations of my supervisee's performance.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I regularly reminded my supervisee of the evaluation standards.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>My supervisee was not surprised with ratings or scores on evaluations.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I consistently monitored my supervisee's progress.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I regularly monitored my supervisee's ethical behaviors.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

What is your gender?
- Male
- Female

What is your age?
What is your race/ethnicity? (Choose all that apply)
- African/Black American
- Asian American/Pacific Islander
- Hispanic/Latino/Latina American
- Native American
- White/European/Caucasian American
- Multiracial
- International
- Others

What is your primary professional identity?
- Licensed Professional Counselor (LPC)
- Licensed Marriage and Family Counselor (LMFC)
- Clinical Social Worker
- School Counselor
- Counseling/Clinical Psychologist
- Psychiatrist
- Ph.D. Student in Counselor Education
- Ph.D. Student in Counseling/Clinical Psychology
- Others (please indicate below)

Approximately, how many years have you provided supervision?
- less than 1 year
- 1 - 5 years
- 6 - 10 years
- 11 - 15 years
- 16 - 20 years
- 21 + years
Appendix C

Chae Optimal Supervision Environment Test (COSET)

The following statements describe some of the ways a supervisor may feel about a supervisee or a supervisory experience. With the same supervisee in mind, to what extent do you agree or disagree with each of the following statements? Please check the column which matches your opinion most closely.

<table>
<thead>
<tr>
<th>Emotional Environment Subscale</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>My supervisee felt “safe” during our supervisory sessions.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>My supervisee interacted with me in a genuine manner.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>Our supervisory relationship was characterized by a sense of mutual trust.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>There was a positive atmosphere during our supervisory sessions.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>My supervisee and I shared mutual respect as part of our supervisory relationship.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Learning Environment Subscale</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I was aware of and sensitive to my supervisee’s training needs.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>I matched my supervision approach to my supervisee’s level of experience.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>I tailored supervision to my supervisee’s level of competence.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>I valued my supervisee’s explanations about clients’ behaviors.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>I modeled appropriate personal and professional boundaries.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Power Environment Subscale</th>
<th>Strongly Disagree</th>
<th>Disagree</th>
<th>Agree</th>
<th>Strongly Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>I acknowledged when my supervisee had made progress towards supervision goals.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>I consistently provided evaluation feedback to my supervisee.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>I was aware of and sensitive to the supervision evaluative process.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>I provided evaluative feedback based on observations of my supervisee’s performance.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
<tr>
<td>I regularly monitored my supervisee’s ethical behaviors.</td>
<td>□</td>
<td>□</td>
<td>□</td>
<td>□</td>
</tr>
</tbody>
</table>


Sterner, W. R. (2009). Influence of the supervisory working alliance on 
supervisee work satisfaction and work-related stress. *Journal of Mental 
Health Counseling, 31*, 249-263.

perspective: The counselor complexity model. *Journal of Counseling 

Stoltenberg, C. D. (2005). Enhancing professional competence through 
developmental approaches to supervision. *American Psychologist, 60*, 
857-864.

An integrated developmental model for supervising counselors and 

Swanson, J. L., & O'Saben, C. L. (1993). Differences in supervisory needs and 
effects of trainee experience, cognitive style, and program 

different supervisory environments and counselor development. *Journal of 
Counseling Psychology, 36*, 336-344.

in psychotherapy supervision: Female trainee perspectives. *Counselling 
and Psychotherapy Research, 7*, 12-18.

Wasserman, J. D., & Bracken, B. A. (2013). Fundamental psychometric 
considerations in assessment. In J. R. Graham & J. A. Naglieri (Eds.),


Vita

Ki Byung Chae

Date of Birth: July 6, 1979
Place of Birth: Junju, South Korea
Education:

2010-2013  The College of William and Mary
            Williamsburg, VA
            Doctor of Philosophy

2006-2009  Liberty University
            Lynchburg, VA
            Master of Art

1999-2006  Chung-Ang University
            Seoul, South Korea
            Bachelor of Arts