Summer 8-21-2022

Exploring the Dimensional Structure of a Measure of Supervision Competence and its Prediction of Trainee Development

Jacqueline Davis
DePaul University, jdavi153@depaul.edu

Exploring the Dimensional Structure of a Measure of Supervision Competence and its Prediction of Trainee Development

A Dissertation

Presented in

Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

By

Jacqueline Olivia Davis-Wright

July 2022

Department of Psychology

College of Science and Health

DePaul University

Chicago, Illinois

Dissertation Committee

Susan D. McMahon, Ph.D., Co-Chairperson

Shona N. Vas, Ph.D., Co-Chairperson

Jocelyn Smith Carter, Ph.D.

Trina Davis Dee, Ph.D.

Yan Li, Ph.D.

Steven A. Miller, Ph.D.

Alexandra Novakovic, Ph.D.

Acknowledgements

I would like to express my sincere appreciation to my advisor and dissertation co-chair, Dr. Susan McMahon: thank you for your support on this project and throughout my graduate school career. Working with you has helped me to grow as a researcher and writer. To my dissertation co-chair and former supervisor, Dr. Shona Vas: I am lucky to have benefited from your genuine commitment to supervision, and I am grateful for your ongoing professional mentorship. Thank you, as well, for allowing me to use data from the UChicago Medicine psychology training programs for this project. To Dr. Steve Miller: thank you for being such a patient and available teacher. Your statistics expertise and support has been integral to this project. To Dr. Jocelyn Carter: thank you for your steady support and guidance throughout my graduate school career. To my family: thank you for your unwavering love and belief in me. Finally, my utmost gratitude and love to my husband, Jonny: thank you for being my rock through it all.

Biography

Jacqueline Davis-Wright (née Davis) was born in Massachusetts, January 10, 1992. She graduated from Sharon High School in Sharon, Massachusetts. She received a Bachelor of Arts degree in Psychology from Boston College in 2014 and a Master of Arts degree in Clinical Psychology from DePaul University in 2018. She is completing her doctoral clinical internship at the West Los Angeles VA Medical Center, where she will also begin a postdoctoral residency in Trauma Psychology in Fall 2022.

Table of Contents

List of Tables

List of Figures

Abstract

Clinical supervision is one of the most important components of a health service psychologist's training. Supervisors ensure the integrity of the supervisee's services to protect the public and act as gatekeepers to the profession. Despite the importance of this professional practice, training in supervision and evaluation of competence to provide supervision received minimal attention until the early 2000s. There is little high-quality research on what makes supervision effective, in part due to few measures assessing supervision competence. A culture shift to competency-based training and education in health service psychology both allows for and requires improved evaluation of supervision. The current study aimed to (a) elucidate the dimensional structure of a measure of supervision competence, and (b) use the results to assess how well supervision competence predicted trainees' development of professional competencies. The study data were collected as part of routine program evaluation within the UChicago Medicine psychology training programs. At the end of each training year from 2015-2020, trainees completed the Psychology Trainee Evaluation of Supervision Competencies (PTESC), a trainee-report measure of supervision competence, and supervisors evaluated interns' acquisition of the nine profession-wide competencies using the Trainee Competency Evaluation. The PTESC has seven domains matching those of the APA's (2014) *Guidelines for Clinical Supervision.* Using 203 responses from 110 trainees, exploratory graph analysis (EGA) was applied using scales for the seven domains to examine the measure's dimensional structure. The EGA revealed a single clique (set of connected nodes), or dimension, of supervision competence. Follow-up confirmatory factor analysis indicated good fit for the single factor model. From the EGA, network scores were generated for the supervision competence clique. Next, cross-classified multilevel modeling was used to assess how well supervision competence predicted

trainee outcomes, as reported by both trainees and supervisors. Supervision competence predicted greater trainee-reported growth across all nine profession-wide competencies but lower ratings of interns' ethics competence by supervisors. Supplemental simple regression models indicated supervision competence predicted interns' growth in professionalism and communication competencies, as reported by supervisors. Trainees' race/ethnicity impacted their report of growth in diversity competence. The findings demonstrate: (1) the utility of the PTESC for assessing supervision competence from the trainee perspective, (2) empirical support for the APA's (2014) seven domains of supervision competence, and (3) that competent supervision enhances trainees' professional competencies, readying them to enter careers in health service psychology. Future directions and implications for research, theory, and practice are discussed.

Introduction

The fields of health service psychology face a conundrum: clinical supervision of trainees is one of the most important components of education and training, yet training in supervision and evaluation of supervision competence received little attention until the early 2000s (Bernard & Goodyear, 2014; Falender & Shafranske, 2021). Supervision is the cornerstone of education and training for health service psychologists (i.e., clinical, counseling, and school psychologists). A clinical supervisor's work serves two primary interconnected purposes: (1) to train the supervisee in provision of clinical services, and (2) to protect the public by ensuring the integrity of the trainee's services (American Psychological Association [APA], 2014; Bernard & Goodyear, 2014; Falender & Shafranske, 2021). Supervisors bear great responsibility: not only do they act as gatekeepers of the profession, but they assume professional and legal liability for their unlicensed supervisees' work with clients (Association of the State Provincial Psychology Board, 2019). The supervisory relationship extends over time and involves evaluation, feedback, and facilitation of the supervisee's skills, knowledge, and self-assessment. Falender and Shafranske (2004) elaborate:

> Supervision provides the structure and framework for learning how to apply knowledge, theory, and clinical procedures to solve human problems. Such experience complements academic and research training, transfers applied knowledge and skills, and establishes competencies in science-informed clinical practice… Supervision provides a relationship in which professional values, commitments, and identity are formed and career goals are formulated. (p.6)

It is no wonder then that supervision has been shown to be among the most important influences on subsequent clinical practice (Lucock et al., 2006; Orlinsky et al., 2005). In a survey of

Australian postgraduate clinical students, participants viewed clinical supervision as the most effective teaching method in their programs (Scott et al., 2011). Additionally, many health service psychologists will go on to supervise others: more than 40% of members of the APA Division of Psychotherapy (Norcross & Rogan, 2013) and the Society of Counseling Psychology (Goodyear et al., 2008) reported currently providing supervision. Despite the importance of this professional practice, training in and evaluation of competence to provide supervision have received minimal attention in the field until recently.

From Decades of Neglect to a "Culture of Competence"

The APA's (2017) *Ethical Principles of Psychologists and Code of Conduct* state that psychologists must practice within their boundaries of competence; therefore, competent practice of supervision is an ethical necessity for supervisors. However, until the past decade or so, many—if not most— supervisors practiced without having had explicit instruction or training in supervision, nor supervision of their supervision. In 2000, Johnson and Stewart found that nearly two thirds of Canadian psychologists who provided supervision had received no formal supervision training. That same year, Scott and colleagues (2000) surveyed training directors of health service psychology doctoral and internship programs. Among academic programs, only 30% required a didactic course/seminar in supervision, and even fewer (23%) required practicum/practice in supervision. Among internships, 35% required a didactic course/seminar in supervision, and 29% required practice in supervision. The bulk of supervision training took place among counseling programs and counseling internship sites (Scott et al., 2000). Therefore, rather than providing supervision based on education and training in this practice, most supervisors used their own supervision experiences and therapeutic orientations to provide supervision. However, over the past two decades or so, a "culture of competence" (Roberts et al.,

2005, p. 356) has emerged within health service psychology training (Callahan & Watkins, 2018).

Competency-based training frameworks grew out of a broader national dialogue around holding training programs and institutions in the health professions accountable for ensuring the quality of their trainees' education (APA, 2014; Falender & Shafranske, 2021). Such frameworks may describe the education, curricula, and training (i.e., "inputs") necessary to become a psychologist and/or the characteristics and capabilities (i.e., "outputs") of a competent psychologist (Falender & Shafranske, 2004). In 2002, the Association of Psychology Postdoctoral and Internship Centers (APPIC) initiated the multinational Competencies Conference: Future Directions in Education and Credentialing in Professional Psychology (hereafter: Competencies Conference; Kaslow et al., 2004). The conference organizers recognized supervision as one of eight core competencies in health service psychology. That same year, the APA's Commission on Accreditation included supervision as a central domain of training (APA, 2002).

Finally in 2014, the APA updated the *Standards of Accreditation for Health Service Psychology* (APA, 2015) to include supervision as one of nine profession-wide competencies (PWCs) in which all graduates of accredited programs must receive training and demonstrate at least a minimal level of achievement consistent with the expectations for independent or entry-level practice. That same year, a Board of Educational Affairs task force convened to develop the first ever *Guidelines for Clinical Supervision in Health Service Psychology* (hereafter: *Guidelines*; APA, 2014) to inform training and practice of competency-based supervision. The task force identified seven broad domains of supervision competence and 29 specific guidelines meant to enhance supervision competence, promote the delivery of quality supervision, and

thereby protect the public. Despite changes in training requirements and guidelines, as well as a growing acknowledgment of supervision as a distinct professional competency, a disconnect remains between the stated importance of supervision and its emphasis within training.

The State of Supervision Training

The integration of supervision instruction and training into programs has been slow. Six years after the Competencies Conference, Crook-Lyon et al. (2008) surveyed 233 health service psychology interns and found that while 72% provided supervision to at least one trainee, only 39% had received supervision training. Comparing these interns' experiences based on training setting, interns in counseling centers reported supervising more trainees, receiving more layered supervision (supervision of supervision), and engaging in more supervision training activities than interns in other settings (Crook-Lyon et al., 2011). Qualitative analyses indicated non-counseling-center interns wanted more supervision training and experiences.

The slow integration of supervision training suggests that it is not seen as a training priority and/or that those tasked with providing this training are not well equipped to do so. Stedman and colleagues (2013) surveyed training directors from 201 APA-accredited internship programs regarding their perceptions of required learning objectives. "Valuing" was assessed by time devoted, rank ordering, and progressive elimination. Their results revealed that training directors generally valued supervision poorly compared to other objectives. More recently, Newman et al. (2021) surveyed course instructors (n = 23) providing supervision training in APA-accredited school psychology programs and conducted qualitative analysis of their syllabi (n = 15). They found that instructors generally had limited training in supervision. Review of the syllabi indicated only some included applied/experiential training in supervision, and most had minimal coverage and/or evaluation of three of the domains of competency-based supervision

identified by the APA's (2014) *Guidelines* (i.e., Diversity, Professionalism, and Problems of Professional Competence; Newman et al., 2021). While this study did not include instructors or syllabi from clinical or counseling programs, the results remain concerning, given that all health service psychology fields abide by the same Standards of Accreditation.

These studies paint a collective picture of the state of supervision training: while more trainees are providing supervision today than they were 20 years ago, the training they receive to do so is still lacking and often not a high priority for training programs and directors. The shift to competency-based training within health service psychology provides an opportunity to ameliorate the historical lack of accountability for supervision-specific training. Simply relying on one's own past experiences of supervision is an ineffective approach: evidence suggests that developing competence in supervision requires training (Lyon et al., 2008; Milne & James, 2002; Milne, Sheikh, et al., 2011). We must be able to identify supervision-specific competencies and evaluate the quality of supervision in order to assess this important PWC.

Supervision Competencies

Many scholars, workgroups, and professional bodies have worked to identify the core competencies of supervision (see Table 1). The Competencies Conference supervision workgroup produced the first consensus statement on supervision competence and a guiding framework for competency-based supervision (Falender et al., 2004). The framework included six areas of knowledge, 12 sets of skills, and 10 values necessary to provide supervision (Falender et al., 2004). The next major advancement in identifying supervision competencies was put forth by the Assessment of Competency Benchmarks Work Group (Fouad et al., 2009). The Benchmarks work group divided competence in supervision into six essential components: expectations and roles, processes and procedures, skills development, awareness of factors

Table 1

Conceptions of Supervision-Specific Competencies

Authors	Supervision Competencies
Competencies Conference Supervision Workgroup (Falender et al., 2004)	**Supraordinate factors:** Recognition that… • "Acquiring supervision competencies is a life-long cumulative, developmental process with levels of proficiency beyond competence… • Attention to diversity in all its forms…relates to every aspect of the supervision process and requires specific competence… • Attention to legal and ethical issues is essential… • Training is influenced by both professional and personal factors… • Necessity that both self- and peer assessment occur regularly across all levels of supervisory development" (p. 778) **Knowledge of…** • Area being supervised • Models theories, modalities, and research on supervision • Professional/supervisee development • Ethics and legal issues specific to supervision • Evaluation, process outcome • (and awareness of) Diversity in all its forms **Skills** • Supervision modalities • Relationship skills • Sensitivity to multiple role with supervisee and ability to perform and balance multiple roles • Provide effective formative and summative feedback • Promote growth and self-assessment in the trainee • Conduct own self-assessment process • Assess the learning needs and developmental level of the supervisee • Encourage and use evaluative feedback from the trainee • Teaching and didactic skills • Set appropriate boundaries and seek consultation when supervisory issues are outside domain of supervision competence • Flexibility • Scientific thinking and translation of scientific findings into practice **Values** • Responsible for client and supervisee • Respectful • Responsible for sensitivity to diversity in all forms • Balance between support and challenging • Empowering • Commitment to lifelong learning and professional growth • Balance between clinical and training needs • Value ethical principles • Commitment to knowing and utilizing available psychological science related to supervision • Commitment to knowing one's limitations
Assessment of Competency Benchmarks Work Group (Fouad et al., 2009)	**Essential components:** • Expectations and roles • Processes and procedures • Skills development • Awareness of factors affecting quality • Participation in supervision • Ethical and legal issues *Note: Includes behavioral anchors for each essential component across three developmental levels*
Guidelines on Clinical Supervision (APA, 2014)	• Supervisor competence • Diversity • Supervisory relationship • Professionalism • Assessment, evaluation, feedback • Problems of professional competence • Ethical, legal, and regulatory considerations *Note: Includes specific guidelines for each domain (29 guidelines total)*

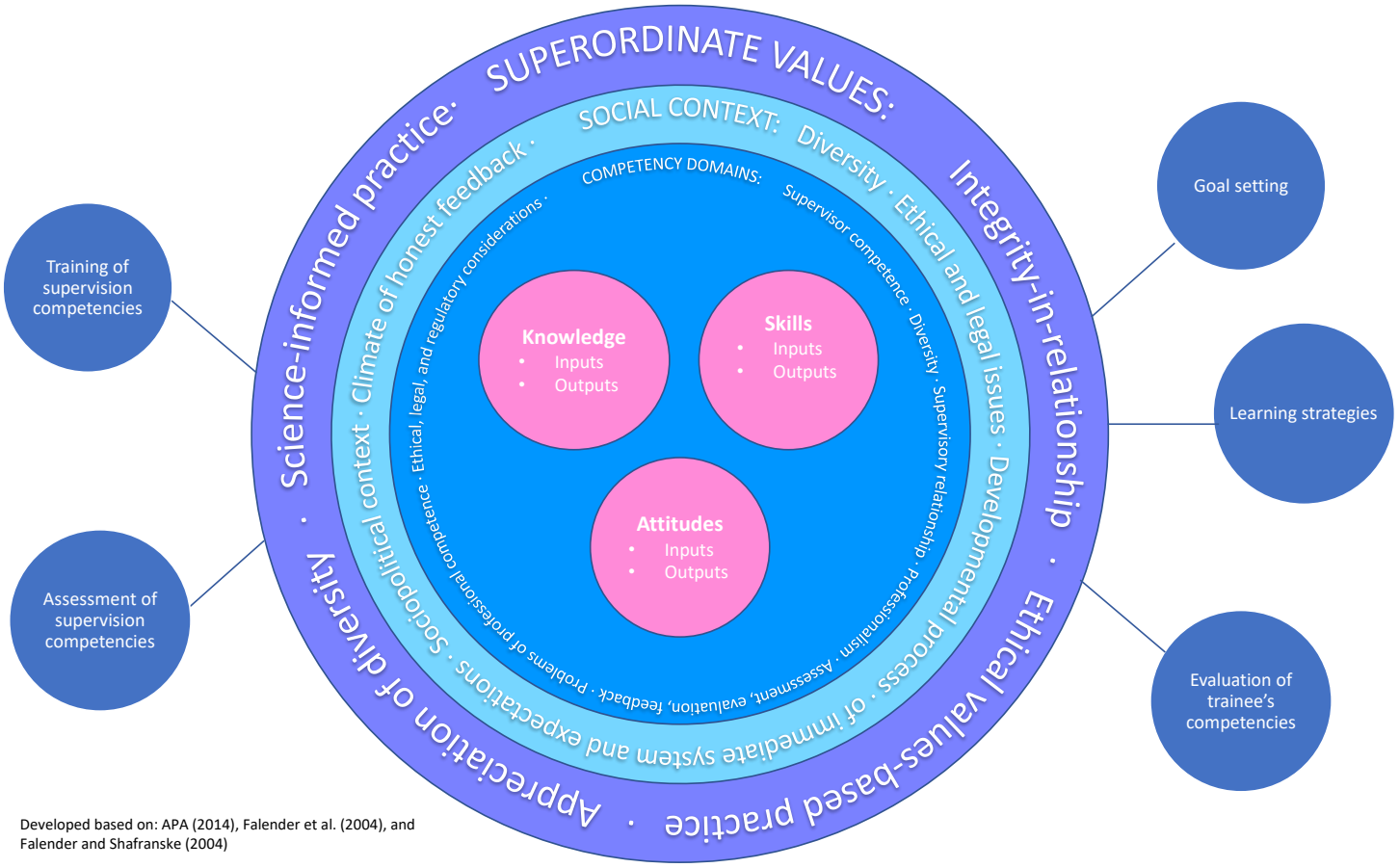
affecting quality, participation in supervision, and ethical and legal issues. They also identified behavioral anchors across these essential components to indicate the threshold for competence in supervision at three different developmental levels (Fouad et al., 2009).

Most recently, the Board of Educational Affairs Task Force on Supervision Guidelines developed the aforementioned *Guidelines* (APA, 2014) to inform training and practice of competency-based supervision. The *Guidelines* are aspirational and non-exhaustive. They were informed by the Competencies Conference supervision work group's framework (Falender et al., 2004) and by guidelines of other regulatory boards and psychological associations worldwide. The *Guidelines* suggest specific professional behaviors across seven competence domains: supervisor competence, diversity, supervisory relationship, professionalism, assessment/evaluation/feedback, problems of professional competence, and ethical, legal, and regulatory considerations (APA, 2014).

These multiple conceptions of supervision competencies can be integrated into a broader framework of competency-based supervision as shown in Figure 1. The framework includes four overarching guiding values (Falender & Shafranske, 2004), which inform supervision: integrity-in-relationship, ethical values-based practice, appreciation of diversity, and science-informed practice. At the next level, factors related to the social context are taken into account, such as an atmosphere of honest feedback and awareness of the sociopolitical climate (Falender & Shafranske, 2004). Within these two layers are the seven competency domains (APA, 2014), which are each comprised of specific knowledge, skills, and attitudes. Specific inputs and/or outputs may be identified to assess whether the competency has been achieved (APA, 2014). Finally, the arms extending from the central components of the framework represent the procedural components. On the right are the three main procedures carried out by the

Figure 1

Framework for Competency-Based Supervision



Developed based on: APA (2014), Falender et al. (2004), and Falender and Shafranske (2004)

supervisor: (1) establishment of supervision goals and objectives, (2) the learning process (i.e., use of educational methods and supervision techniques; Milne & James, 2002), and (3) evaluation of the trainee. On the left are the preparatory and monitoring procedures focused on the supervisor: training and evaluation of supervision competencies.

This study directly assessed the APA's (2014) supervision competence domains represented in the innermost blue circle of Figure 1; however, these domains are intricately related to, and even overlap with, the superordinate values (e.g., ethical values-based practice) and procedures (e.g., evaluation) of supervision. Although these *Guidelines* were thoughtfully considered based on existing supervision theory and best practices, they have yet to be empirically validated for two main reasons: methodological weaknesses of the supervision literature generally and the lack of measures focused on supervision competence.

Supervision: The Good, the Bad, and the Unknowns

There have been growing calls within health service psychology for greater accountability for the quality of supervision (e.g., Watkins, 2011; Ellis, 2010). Unfortunately, methodological issues within the extant literature limit empirical findings on the factors contributing to high quality supervision. Common weaknesses of the research include inadequate power, poor methodology, Type I and II errors, an absence of outcomes research, lack of theoretical foundations, ambiguous hypotheses, and lack of psychometrically sound measures (Bambling et al., 2006; Ellis & Ladany, 1997). Historically, reviews of the supervision literature have found its overall quality to be poor: Ellis and colleagues (1996) and Ellis and Ladany (1997) conducted two seminal reviews of the supervision literature using similar search procedures. Ellis et al.'s (1996) review of 144 supervision studies found that 80% or more had inflated Type I or Type II error rates or unreliable dependent or independent measures. Other

common methodological threats included nonrandom or nonrepresentative samples (42.36%); mismatch of purpose, hypothesis, design method, and analyses (26.39%); and violated assumptions of statistics (14.58%). Ellis and Ladany's (1997) review (n = 104 studies) similarly concluded that the supervision literature lacked conceptual and methodological rigor. Subsequent reviews (e.g., Freitas, 2002; Watkins, 2011) have noted some but little improvement in the state of the literature.

To improve the quality of supervision literature, Freitas (2002) recommended researchers (a) recruit therapist participants who are not trainees and have similar training backgrounds and levels, (b) use fairly uniform clients, (c) randomly assign clients to supervisees, and (d) obtain client outcome measures from multiple reporters. These recommendations are logistically difficult, if not impossible, to implement in a standard training clinic and therefore raise questions regarding ecological validity and generalizability. Nonetheless, the importance of clinical supervision warrants innovation in research to establish a stronger evidence base. Two particular weaknesses in the literature are: 1) the minimal attention paid to evaluating supervisor competence, and 2) the lack of high-quality studies on how supervision impacts supervisee competency development (Falender & Shafranske, 2017).

Good Supervision

Although not the same as *competent* supervision, many researchers have sought to answer the question, "What makes for 'good' supervision?" over the past few decades. The one variable that seems to consistently emerge from the literature is the supervisory relationship or alliance. Given the inherent relational nature of supervision, the supervisory relationship has been posited as the foundation of effective supervision (Bernard & Goodyear, 2014; Ellis, 2010). A supportive supervisory relationship is characterized by empathy, warmth, a sense of

teamwork, validation, approachability, attentiveness, respect for autonomy, a strengths-based approach, a nonjudgmental stance, appropriate self-disclosure, and balanced constructive feedback (Falender & Shafranske, 2021; Watkins, 2017). The closely related concept of the supervisory alliance adds to this relational bond an agreement on the tasks and goals of supervision (Bordin, 1983).

In a review of 40 studies on the supervisory alliance published between 1990-2013, Watkins (2014) summarized that a strong alliance was related to: greater trainee self-efficacy, wellbeing, willingness to self-disclose, and availability of coping resources; perceived effectiveness of supervision; perceived ethical behavior of the supervisor; more discussions of culture in supervision; and more frequent but appropriate supervisor self-disclosures. On the contrary, a weak alliance was related to greater trainee stress and burnout, as well as more frequent negative supervision events. Watkins (2014) noted that the literature provided strong clinical support but weaker empirical support for the supervisory alliance (e.g., there was only one randomized controlled trial among the 40 studies). Studies published since Watkins's (2014) review continue to demonstrate the positive influence of a strong supervisory alliance on trainee disclosures (e.g., Gibson et al., 2019; Hutman & Ellis, 2020; Mehr & Daltry, 2014; Mehr et al., 2015). A more recent meta-analysis by Park and colleagues (2019) examined the effect of the supervisory alliance on supervision outcomes in 27 articles, dissertations, and theses published between 1990-2018. Results indicated the supervisory alliance was positively related to supervision outcomes, including supervisee perception of their relationship with the client.

Another factor proposed to influence the quality of supervision is the supervision interventions used (Watkins, 2017), or perhaps or the range of interventions (Milne & James, 2002). Milne and James (2002) developed the "Teachers PETS" (Process Evaluation of Teaching

and Supervision) observational instrument to code interactions between supervisors and supervisees. Their coding revealed supervision was more effective in guiding trainees' experiential learning when the supervisor used a greater range of supervision techniques (e.g., managing, listening, challenging, supporting, informing): supervisees' behaviors responded accordingly (i.e., they demonstrated more conceptualization, experiencing, and experimenting), and they were more satisfied with supervision (Milne & James, 2002). In a qualitative review by Bradley and Becker (2021), the authors found that corrective feedback, discussing intervention, and role play were the most common supervision practices for promoting supervisee formative outcomes (e.g., skill development). One other practice of note is discussing culture: such discussions in supervision appear to positively influence trainee satisfaction with supervision and the supervisory alliance (Gatmon et al., 2001; Phillips et al., 2017; Soheilian et al., 2014). Evidently, there is still much to learn about what makes for good, effective supervision.

Bad Supervision

Unfortunately, more may be known about what "bad" (e.g., harmful or inadequate) supervision looks like. Magnuson et al. (2000) interviewed 11 experienced counselors about their supervision experiences to better understand and characterize "lousy" supervision. Conventional coding elucidated six principles of lousy supervision: the supervision was unbalanced (too much or too little of the elements of supervision), developmentally inappropriate, intolerant of differences (i.e., inflexible), poor model of professional attributes, untrained, and professionally apathetic. Ladany et al. (2013) surveyed 128 trainees regarding effective and ineffective behaviors of supervisors. Trainees reported that depreciating of supervision, ineffective client conceptualization/treatment, and weakening of the supervisory relationship made supervision less effective. Narratives of harmful supervision (Ellis, 2017; McNamara et al., 2017) describe a

wide range of problematic behaviors, from inattentiveness and lack of investment/involvement in the supervision to intimidating, sexually coercive, and physically abusive behaviors. Trainees of diverse identities (e.g., gender, sexuality, and race/ethnicity) have also reported instances of insensitivity, microaggressions, and harassment in supervision (Bautista-Biddle et al., 2021).

Lousy, harmful, or inadequate supervision may be more common than previously thought. Ellis and colleagues (Ellis, 2010; Ellis, Siembor, et al., 2008) surveyed 363 supervisees to explore the prevalence of inadequate and harmful supervision. The researchers were alarmed to find that 50% of trainees were currently receiving inadequate supervision, and 36% were currently receiving harmful supervision. Additionally, 75% and 51% of trainees, respectively, reported receiving inadequate or harmful supervision at some point in their training. Equally concerning, one third of trainees who reported receiving inadequate supervision perceived that the supervision was "moderately" to "totally" harmful to their clients. The prevalence of "bad" supervision and its potential to harm both trainees and/or patients underscore the need for accountability in evaluating supervision and those whom it impacts.

What Impact Does Supervision Have?

While trainees may perceive harmful supervision as negatively impacting their clients (Ellis, 2010; Ellis, Siembor, et al., 2008), the flip side of this issue is more promising: supervisees' satisfaction with supervision is significantly positively correlated with its perceived impact on their clinical practice (Kavanagh et al., 2003). The impact of supervision is generally discussed in terms of trainee and/or client/patient outcomes. More evidence exists for the influence of supervision on trainees than patients. While many scholars have critiqued the supervision literature for this reason (e.g., Ellis & Ladany, 1997; Watkins, 2011), Reiser and Milne (2014) argue that too much emphasis has been placed on whether supervision improves

patient outcomes: ensuring safe practice is the foremost goal of supervision, which is ultimately a measure of trainees' behavior. Studying trainee outcomes, therefore, remains essential.

Trainee Outcomes. Much of the research on supervision outcomes fails to differentiate between satisfaction with supervision and effectiveness of supervision (Falender & Shafranske, 2021; Gonsalvez et al., 2017), yet supervision is most commonly assessed via trainee satisfaction (Milne, 2009). Satisfaction may assess the quality of the supervisory relationship, but it does not necessarily equate to growth as a trainee or effectiveness with clients. Wheeler and Richards (2007) conducted a systematic review of literature published 1980-2006 examining the impact of supervision on therapists/counselors and their clients. Notably, they excluded studies looking at self-reported satisfaction with supervision. Their search yielded 18 studies: eight were quantitative, three were qualitative, and seven used mixed methods. Each study was rated on its (a) methodological rigor and (b) overall quality using a scale of 1-5. Wheeler and Richards (2007) concluded that the majority of the studies lacked methodological rigor, but supervision appeared to have positive effects on supervisees' self-awareness, skills, and self-efficacy. The review therefore suggested supervision may have a positive impact on some supervisee outcomes, but the quality of the studies precluded the authors' ability to draw any strong conclusions regarding client outcomes. This inconclusive summary echoed those of many other review articles exploring the influence of supervision on patient outcomes (e.g., Ellis et al., 1996; Ellis & Ladany, 1997; Freitas, 2002; Holloway & Neufeldt, 1995; Milne et al., 2008; Roth et al., 2010; Watkins, 2011).

In an empirical review of 24 supervision studies, Milne and colleagues (2008) found the most frequently reported trainee outcomes were experiencing (attitude change, affective awareness, motivation), reflection (self-awareness, -evaluation, and -monitoring), planning,

conceptualizing, experimenting, and general learning. In a relatively more recent (but not systematic) review of the literature on supervision outcomes, Bambling (2014) summarized:

> There is sufficient evidence to conclude supervision creates a variety of positive outcomes... Supervision may enhance supervisee self-efficacy, knowledge, and skills, at least in the training setting. Most encouragingly, there is evidence that supervision might also improve the quality of client work and enhance treatment outcomes for clients. Process factors, such as supervisory alliance, are important to ensure the quality of supervision and the achievement of learning goals and clinical outcomes. There is insufficient data as yet to explain the mechanisms by which supervision achieves these outcomes. (p. 453)

Similarly, in a review by Callahan and Watkins (2018), the authors conclude that supervision positively impacts trainee skill acquisition, treatment knowledge, self-awareness, self-efficacy, and working alliance with clients. Encouragingly, supervision that improves therapist adherence to a treatment protocol (e.g., Schoenwald et al., 2009) can be expected to improve client outcomes, as well (Reiser & Milne, 2014).

Patient Outcomes. Regarding supervision's influence on the patient/client, empirical findings are limited and mixed. Watkins (2011) conducted a review of 30 years of literature (1981-2011) to find studies looking at patient outcomes of clinical supervision. Although 18 studies were identified, only three were found to meet reasonable standards of methodological rigor (Bambling et al 2006; Bradshaw et al., 2007; White & Winstanley, 2010). Of these three studies, two found positive effects of clinical supervision, including greater symptom improvement or reduction of symptoms (Bambling et al., 2006; Bradshaw et al., 2007), as well

as improved working alliance and lower rates of attrition (Bambling et al., 2006). White and Winstanley (2010), however, found no effect of supervision in a sample of mental health nurses.

In a large review of the literature on individual supervision from 1994-2012 (n = 233 articles), Inman and colleagues (2014) concluded that supervision appears to have a significant influence on client symptom reduction and treatment retention. The size of this effect, however, is unclear. Callahan et al. (2009) found that supervisors had a moderate sized effect on client outcome, and the supervisor accounted for 16.4% of variance in patient change scores (change in distress from first to last session). Wrape et al. (2015) replicated this finding in a larger sample. Rousmaniere et al. (2016), however, found essentially no effect of supervision across assorted mental health disciplines, with less than 1% of variance in client outcomes associated with supervisors. Taking these findings together, Callahan and Watkins (2018) suggested supervision may account for 1-16% of the variance in client outcomes. The variability in the size of the effect of supervision may depend upon the supervisor's competence (Callahan & Watkins, 2018), providing yet another reason for improved assessment of supervision competence.

Evaluating Supervision

The gaps and methodological issues within the existing supervision literature underscore the need for improved evaluation of supervision competence and outcomes; however, little attention has been paid until recently to the idea of assessing supervision competence (Bernard & Goodyear, 2014; Falender et al., 2014). The Competencies Conference supervision work group (Falender et al., 2004) provided recommendations for assessment of supervision competencies, including documented supervisee feedback and assessment of supervisee outcomes. While there are many instruments for assessing supervision, there is a dearth of measures with demonstrated

psychometric validity and reliability (Ellis, D'Iuso, et al., 2008), particularly when it comes to assessing supervision competence.

A review by Ellis, D'Iuso, and Ladany (2008) of clinical supervision measures revealed a near complete lack of psychometrically valid or reliable instruments. The authors conducted a literature search for articles published 1995-2007 focused on assessment of clinical supervision and the development of a measure and its psychometric properties. They identified six viable articles and evaluated their scientific rigor and the psychometric properties of the measures within. The sole measure recommended for use was the Evaluation Process Within Supervision (EPSI; Lehrman-Waterman & Ladany, 2001). The EPSI assesses trainees' perceptions of the effectiveness of goal setting and feedback in supervision, but it is not a broad measure of competence in supervision.

Wheeler and Barkham (2014) combed the supervision literature starting in 1980 to identify supervision instruments and develop a core evaluation battery for supervision. Their search yielded 150 measures, of which they were able to obtain copies of 67. The authors developed a rating system for evaluating and scoring the measures on seven criteria (e.g., pan-theoretical, relatively short, face validity). They ultimately included five measures in the battery. These measures collect data on the supervisory relationship/alliance, supervisee and supervisor biographical characteristics, helpful aspects of supervision, and supervisee role conflict or role ambiguity. They also recommended two additional measures evaluating the supervisory relationship/alliance. None of these seven measures assess competence in supervision. Competence-specific measures have only emerged within the field quite recently.

Existing Competence Measures

There are five known measures of supervision competence. The Supervisor Competency Self-Assessment Checklist (Falender et al., 2016) is a 28-item, self-report checklist supervisors can use to rate their own competencies across the APA's seven domains of competent supervision and identify areas for improvement. While this measure may be a helpful self-assessment tool, it has yet to gain widespread use or be empirically evaluated, and its ratings are susceptible to self-report bias.

The Supervision: Adherence and Guidance Evaluation (SAGE) is an observational instrument for evaluating competence in cognitive behavioral therapy (CBT) supervision (Milne, Reiser, et al., 2011). In particular, the instrument is used for assessing the supervisor's facilitation of supervisee learning. Items are focused on supervisor behaviors such as agenda setting, discussing, evaluating, formulating, listening, prompting, and teaching. The SAGE has demonstrated acceptable face, predictive, and discriminant validity, as well as interrater reliability (Milne, Reiser, et al., 2011). Follow-up research and principal components analysis revealed a two-factor structure for the SAGE, and the instrument was shortened from 23 to 14 items (Reiser et al., 2018). The trimmed instrument showed high internal reliability. Nonetheless, the SAGE has several limitations: it is specific to CBT, it is focused solely on the supervisor's facilitation of learning, and its administration method (direct observation) presents logistical challenges (e.g., administration training, the time needed to review supervision sessions).

The Supervision Evaluation and Supervisory Competence Scale (SE-SC) is a 31-item, supervisee-report measure of supervisory competence (Gonsalvez et al., 2017). The measure includes six "overall evaluation of supervision" items and 25 supervisor competency items. Items were developed from Gonsalvez and colleagues' (2002) Objectives Approach to

Supervision, an early model of competency-based supervision. Using responses of 142 supervisees and 22 items of the scale for which there were sufficient data, hierarchical cluster analysis revealed a six-cluster solution with good internal and test-retest reliability, as well as concurrent and convergent validity. The clusters collectively predicted 85% of variance in supervisee-reported satisfaction and effectiveness. The authors admit that while the 22 items included in the analyses cover most supervision competencies, the supervisor's competence in dealing with ethical, legal, and multicultural issues was omitted, as was competence in providing summative feedback.

Finally, the Generic Supervision Assessment Tool (GSAT; Hamilton et al., 2022) is the newest measure of supervision competence. This tool is non-discipline specific. Items were generated by a diverse group of stakeholders (e.g., supervisors, clinical educators, academics, and supervisees in psychology, nursing, occupational therapy, counseling, and social work) in Australia and New Zealand. Supervisor-rated (GSAT-SR) and supervisee-rated (GSAT-SE) versions of the measure were developed: the GSAT-SR has four factors and 26 items, while the GSAT-SE has two factors and 21 items. Exploratory and confirmatory factor analysis were used to refine and validate these measures, which have demonstrated good internal, convergent, and face validity.

The SE-SC (Gonsalvez et al., 2017) and GSAT (Hamilton et al., 2022) are particularly promising measures of supervision competence; however, neither is grounded in the APA's seven domains of competency-based supervision. Additionally, results of the psychometric validation analyses of the SE-SC omitted several important competencies (used 22 of 31 items), and the GSAT is not a psychology-specific measure. This study is the first to empirically

evaluate the APA (2014) *Guidelines* using a new and comprehensive measure of supervision competence: the Psychology Trainee Evaluation of Supervision Competencies (PTESC).

Rationale

Supervision is an extremely important component of training in health service psychology. Supervisors safeguard both the public and the profession (Bernard & Goodyear, 2014), and they have substantial influence on trainees' subsequent clinical practice (Lucock et al., 2006; Orlinsky et al., 2005). Many, if not most, mental health professionals will provide supervision at some point in their careers (Goodyear et al., 2008; Norcross & Rogan, 2013). However, psychologists have historically received very little training on how to be a supervisor (e.g., Johnson & Stewart, 2000; Scott et al., 2000), and harmful and inadequate supervision is alarmingly common (Ellis, 2010; Ellis, Siembor, et al., 2008). High-quality empirical literature on (a) what makes supervision effective, and (b) the impacts of supervision on trainees and patients remains limited (Callahan & Watkins, 2018; Watkins, 2011, 2014; Wheeler & Richards, 2007). These issues are due, in part, to a lack of psychometrically sound measures for assessing supervision competence (Ellis, D'Iuso, et al., 2008; Wheeler & Barkham, 2014). Such measures are only just recently emerging and have yet to gain widespread use.

Norms around training in supervision have changed over the past 20 years. Education and training programs in health service psychology are adopting competency-based approaches, which both provide and require greater accountability for quality training (APA, 2014; Falender & Shafranske, 2021). Supervision is one of nine PWCs expected of graduates of health service psychology programs (APA, 2015). Competency-based supervision represents a perspective shift from previous approaches: it emphasizes not only the trainee's competence in developing the necessary knowledge, skills, and attitudes to enter into the profession, but also the supervisor's

competence in providing supervision. However, the APA's (2014) *Guidelines for Clinical Supervision* have yet to be empirically tested, and measures of competence in supervision remain few and limited. This state of affairs presents a major problem, given the centrality of supervision in health service psychology training. The PTESC (Vas et al., 2021) was developed to meet the need for a comprehensive measure of supervision competence and to empirically assess the APA's (2014) *Guidelines* and seven domains of competent supervision.

The current study applied exploratory graph analysis to the PTESC to: (a) reveal the dimensional structure of the measure and supervision competence more broadly, (b) assess how well the dimensional structure aligns with the APA's (2014) seven domains of competent supervision, and (c) examine the measure's value in predicting trainees' acquisition of professional competencies. This information will add to the literature by offering empirical support for the APA (2014) *Guidelines*, clarifying factors associated with competent supervision, and evaluating their influence on trainee competencies, going beyond simplistic assessment of satisfaction with supervision.

Statement of Research Questions

Research Question 1. What is the dimensional structure of the Psychology Trainee Evaluation of Supervision Competencies (PTESC) measure?

Research Question 2. To what extent are the dimensions related to one another?

Research Question 3. How well do the dimensions of supervision competence predict trainee acquisition of the nine profession-wide competencies (PWCs; i.e., research and scholarship; ethical and legal standards; individual and cultural diversity; professional values, attitudes, and behavior; communication and interpersonal skills; assessment; intervention; supervision; and consultation and interprofessional/interdisciplinary collaboration)?

a. How well do the dimensions predict trainees' self-reported acquisition of the PWCs?
b. How well do the dimensions predict trainees' acquisition of the PWCs as reported by supervisors?

Methods

Participants

The sample consisted of 203 responses to the PTESC by 110 trainees from the health service psychology training programs of the University of Chicago Medicine (UCM). More specifically, 85 participants were externs (i.e., pre-internship psychology doctoral students), 22 were interns (i.e., psychology doctoral students completing their capstone training year), and 3 were postdoctoral fellows (i.e., pre-licensure PhDs). Sample demographics were not collected as part of the program evaluation data being used in this study. However, internal data from the internship program and first-hand knowledge from program faculty allowed for identification of gender and race/ethnicity for 84.09% of the sample. Of this subset, trainees were 77.17% White, 11.96% Asian, 5.43% Latinx, 4.35% Black or African American, and 1.01% multiethnic. Participants were 89.25% female.

Materials

The Psychology Trainee Evaluation of Supervision Competencies (PTESC)

The Psychology Trainee Evaluation of Supervision Competencies (PTESC; Vas et al., 2021; Appendix B) is a 105-item measure of competence in supervision as rated by trainees. The PTESC was developed in 2015 to address the need for a measure of supervision competence and for program evaluation within the UCM psychology training programs. The PTESC has seven domains matching those of the APA's (2014) *Guidelines:* Supervisor Competence (31 items); Diversity (15 items); Supervisory Relationship (12 items); Professionalism (8 items);

Assessment, Evaluation, and Feedback (15 items); Trainee Remediation and Management of Problems of Professional Competence (7 items); and Ethical, Legal, and Regulatory Considerations (8 items). Each domain contains one overarching "goal item" asking how frequently the supervisor displayed behavior aligned with the goal of this domain. The Supervisor Competence domain is the one exception, as it includes two goal items: one focused on competence in the clinical services being provided, and one focused on competence in provision of supervision (e.g., seeking supervision competence through education and training; knowledge of supervision literature). After the goal item, each domain includes items asking about the frequency of more specific, relevant behaviors. For instance, the goal item for Domain C: Supervisory Relationship is: "Creates a supervisory relationship that facilitates effective clinical supervision." Subsequent items include: "Demonstrates respect for trainees, patients, and colleagues" and "Promotes growth and self-assessment in trainee." The 96 items across these seven domains were drawn from specific guidelines for competency-based clinical supervision put forth by the APA (2014). They are each rated on a 5-point Likert scale from *Behavior Never Displayed/Observed* to *Behavior (Almost) Always Displayed.* A sixth *Not Applicable* option is also provided.

Nine additional "growth" items at the end of the PTESC ask the trainee to what degree the supervisor facilitated their acquisition of each of the profession-wide competencies (PWCs). For instance, the Diversity Growth item reads: "My training experience with this supervisor facilitated the acquisition of competency in individual and cultural diversity." These nine items are rated on a 5-point Likert scale from *Strongly Disagree* to *Strongly Agree*.

Each trainee completed PTESC evaluations for an average of 2 supervisors (SD = 1.28, range = 1-5). Each supervisor was evaluated by an average of 6.15 trainees (SD = 6.89, range =

1-30). The average time point at which the PTESC was completed was 11.86 months (SD = 1.00, range = 9-15).

The UChicago Medicine Psychology Trainee Competency Evaluation (TCE)

The UCM Psychology Trainee Competency Evaluation (TCE; Appendix C) is a 199-item, supervisor-report measure used internally within the UCM psychology training programs to provide trainees (and their respective graduate training programs) with formal summative feedback on their performance. The TCE has nine domains aligning with the nine PWCs for health service psychologists: Science, Research, and Evaluation (18 items); Ethical and Legal Standards (11 items); Individual and Cultural Diversity (13 items); Professional Values, Attitudes, and Behaviors (25 items); Communication and Interpersonal Skills (12 items); Psychological Assessment and Diagnosis (54 items); Psychotherapeutic Intervention (43 items); Supervision, Education, and Training (12 items); and Consultation and Interprofessional/Interdisciplinary Collaboration (11 items). Each domain contains one or more "goal" items (range of 1-6 goal items per domain) followed by relevant "specific objective" items (range of 3-10 items per goal item). Each item is rated on a 1-5 scale of competence: 1 = *Needs Remediation*, 2 = *Entry Level (Continued intensive supervision is needed)*, 3 = *Intermediate (Should remain a focus of supervision)*, 4 = *High Intermediate (Occasional supervision needed)*, and 5 = *Advanced (Skills comparable to autonomous practice at the licensure level).* There is also a sixth *Not Applicable or Not Observed During this Training Experience* response option. TCE data were available only for the intern subset of the sample. Interns were evaluated by an average of 4 supervisors using the TCE (SD = 1.00, range = 1-5). Interns must receive an average score of 4 on each PWC by the end of the internship year to successfully complete the program.

Procedure

Data for the study came from the UCM psychology training programs. UCM's Department of Psychiatry and Behavioral Neuroscience offers externship, internship, and postdoctoral fellowship training programs. There are a wide variety of externship programs for health service psychology doctoral students in psychotherapy and assessment with youth and adult patients. The department's APA-accredited psychology internship program typically accepts 5 interns per year. The availability of postdoctoral fellowship positions varies from year to year based on funding. From fall 2015 to spring 2020, the UCM psychology programs trained 233 externs, 24 interns, and 14 fellows.

The study data were collected as part of routine program evaluation (exempt from UCM IRB review) of the externship, internship, and fellowship programs. The current study was deemed exempt from DePaul University IRB review as secondary research in which participants are not readily identifiable and will not be recontacted or reidentified. In 2015, UCM trainees were asked to evaluate each of their supervisors via the PTESC measure at the 3-, 6-, 9-, and 12-month marks of the training year. From 2016-2020, trainees were asked to complete the measure only at 6 and 12 months. The department's Director of Clinical Psychology Training ("Training Director") emailed a link to all trainees to complete the PTESC online at their convenience. From years 2015 to 2018, the survey asked for the trainee's name but could also be completely anonymously. From 2019 onward, the survey instructed the trainee to create a unique and confidential identifier for themselves rather than including their name.

Completing the PTESC was optional for externs and postdocs. Due to the internship program's accreditation requirements, interns were required to complete the PTESC and to provide their names in completing the measure. Trainees were notified that their responses would

be used for program evaluation and quality improvement. The Training Director's email asked trainees to be candid in their evaluations, while acknowledging the inherent power differential involved in supervision. The instructions noted that responses would remain confidential and not be shared directly with supervisors without the trainee's permission; data would only be shared with supervisors in anonymous and aggregated form. The one exception to this was the Training Director, who collected the evaluations and was also a supervisor.

Across the 2015-2020 training years, supervisors evaluated interns' PWC acquisition using the TCE at the midpoint (6 months) and endpoint (12 months) of the training year. The TCEs were completed in hard copy format with the intern and supervisor name identified. Complete TCE responses were not available; the Training Director provided an abbreviated, deidentified dataset with mean ratings for each of the PWC domains, rather than individual responses to each item. The abbreviated TCE dataset allowed for removal of identifiers from the data and transfer of the data to digital format while retaining the relevant variables of interest.

Data Preparation

The initial dataset included a total of 559 PTESC evaluations completed between fall 2015 and spring 2020. For the purposes of these analyses, the dataset was trimmed to include only endpoint data, defined as 12 +/- 3 months. If a trainee evaluated a supervisor more than once within the designated endpoint timeframe (e.g., at 9 and 12 months), the later evaluation was kept, and the earlier one was removed. Use of a single timepoint was necessary to minimize issues of dependency inherent in repeated measures analysis and to obtain an interpretable dimensional structure of the PTESC.

The dataset was further trimmed in the process of data cleaning. Two blank PTESC evaluations were removed. Four trainees submitted duplicate or nearly identical evaluations of

the same supervisor at or near the same timepoint (i.e., one month apart or less). Duplicate entries were removed. In the case of minor discrepancies (i.e., a 1-point difference in a rating) on nearly identical evaluations, means of the item responses were calculated and retained.

Based on the final PTESC dataset, matched evaluations were available for 87 intern-supervisor dyads (e.g., Intern A completed the PTESC about Supervisor B, *and* Supervisor B completed the TCE for Intern A). These matched pairings were added to the dataset. There were four instances where a supervisor completed two TCEs for an intern for different clinics/rotations. In these instances, the mean of the PWC ratings was used. The final dataset comprised 203 responses to the PTESC by 110 trainees evaluating 33 unique supervisors, plus 87 matched TCE responses.

Statistical Analyses

Analyzing the Dimensional Structure of the PTESC

Multilevel Exploratory Factor Analysis (ML-EFA). Factor analysis is a statistical method which uses structure-analyzing procedures to reveal the relationships between observed variables, and to then group related subsets of variables into latent "dimensions" or "factors" (Pett et al., 2003). Multilevel exploratory factor analysis (ML-EFA) is used to explore the underlying factor structure of a set of variables with nested data (e.g., trainees nested within supervisors) when one does not yet know the number of factors necessary to explain the relationships among the variables (Kim et al., 2016). Use of ML-EFA was planned to examine the dimensional structure of the PTESC; however, it could not be conducted due to lack of variation on individual items within many of the clusters (i.e., low intraclass correlation coefficients, meaning different trainees' evaluations of the same supervisor showed little variability, and/or there were not enough evaluations of the supervisor to achieve sufficient

variability). Exploratory graph analysis (EGA) was used as an alternative statistical method for elucidating the dimensional structure of the dataset.

Exploratory Graph Analysis (EGA). EGA (Golino & Epskamp, 2017) is a type of psychometric network modeling. Network approaches are quickly gaining popularity in psychology research (Christensen et al., 2019). At the most basic level, network models consist of nodes, edges, and communities (Christensen et al., 2019). Nodes are the observed variables. Edges, visually represented as lines connecting nodes, represent the partial correlations given all other nodes. Communities, the statistical equivalent of latent factors, are sets of connected nodes (Golino & Epskamp, 2017). Communities may also be referred to as *cliques* when all of the nodes in the network section are connected or *clusters* when most of the nodes are connected. EGA estimates the correlation matrix of the observed variables and then applies a Gaussian Graphical Model (computed using the graphical least absolute shrinkage and selection operator [glasso]) to obtain the inverse covariance matrix (Christensen et al., 2019; Golino & Epskamp, 2017). The walktrap community detection algorithm is then applied to identify dense subgraphs (i.e., communities) of the partial correlation matrix. The walktrap algorithm does this via use of "random walks," which are jumps from one node to another where each node is alternatingly used as the starting point (Golino & Epskamp, 2017). The random walks cross neighboring edges, and greater edge weights (i.e., higher partial correlation coefficients) are more likely to be crossed. Communities thereby form based on the nodes' proportions of densely versus sparsely connected edges. The number of communities detected is equivalent to the number of dimensions. This procedure yields high accuracy in estimating the dimensional structure of a dataset and avoids overfitting because it shrinks smaller partial correlation coefficients (estimates them to be zero), which aids interpretability (Christensen et al., 2019; Golino & Epskamp, 2017).

EGA has been shown to be effective in discovering new dimensions of constructs, as well as replicating factor analytic findings (Bell & O'Driscoll, 2018; Christensen et al., 2018). EGA performs as well as, if not more effectively, than principal components analysis, common factor analysis, and parallel analysis with both real-world and simulated datasets (Christensen et al., 2018; Golino & Demetriou, 2017; Golino & Epskamp, 2017). Notably for this study, parallel analysis tends to underestimate the number of factors when sample sizes are small (< 500) or when correlations between factors are high (Crawford et al., 2010; Green et al., 2016; Keith et al., 2016; Ruscio & Roche, 2002). In such cases, EGA may be a better statistical option (Golino & Epskamp, 2017). Furthermore, communities discovered via EGA are deterministic, requiring no direction from the researcher—a major contrast from the relatively subjective factor extraction decision-making process in factor analysis (Christensen et al., 2019; Golino et al., 2020). EGA is also unique in its production of a visual network plot. Both of these characteristics of EGA facilitate interpretability of the results. EGA, therefore, was selected as an alternative statistical analysis to the planned ML-EFA for its accuracy, interpretability, and ability to handle smaller sample sizes.

EGA was applied in RStudio (version 1.4.1106; R Core Team, 2021) using the *EGAnet* package (version 0.9.8; Golino & Christensen, 2021), which uses the *qgraph* package (version 1.6.9; Epskamp et al., 2012) for visualizations of networks. EGA was initially attempted using all the specific behavioral items of the PTESC; however, the results indicated the correlation matrix was not positive definite, and the results were unreliable. Therefore, scales were created for the seven PTESC domains by taking the mean of each domain's specific behavioral items. Compared to using individual items, scales are more reliable and more likely to be normally distributed, and models using scales (as opposed to items) have fewer parameter estimates and

sources of sampling error (Little et al., 2013). EGA was applied with the seven scales to examine the dimensional structure of the PTESC. Bootstrapping was then used to replicate the EGA 500 times (Davidson & MacKinnon, 2001; Pattengale et al., 2010) to determine the frequency with which the initial solution would be repeated.

Confirmatory Factor Analysis (CFA). Confirmatory factor analysis (CFA) was subsequently applied to evaluate the EGA model. The weighted least square mean and variance (WLSMV) estimator was used; this is the appropriate estimator for ordinal data, and RStudio treated the dataset as ordinal given the small number of measure response options. Fit statistics for the model were generated using the *lavaan* package (version 0.6-8; Rosseel, 2012).

Cross-Classified Multilevel Modeling (CCMLM)

Cross-classified multilevel modeling was used to assess how well the supervision competence clique obtained from the EGA (see Results) predicted trainee acquisition of the nine profession-wide competencies (PWCs). Multilevel modeling is useful when working with nested data structures because it allows for modeling of within- and between-group variance and examination of "the influence of higher level units on lower level outcomes while maintaining the appropriate level of analysis" (Hofmann, 1997, p. 726). Cross-classified multilevel modeling (CCMLM) more specifically is used when the lower level unit belongs to more than one higher level unit (Leckie, 2013). Had each trainee only worked with one supervisor, traditional multilevel modeling would have been appropriate; however, in this dataset, trainees were supervised by and evaluated more than one supervisor. Therefore, CCMLM was planned to assess how well the clique would predict trainees' acquisition of PWCs while simultaneously accounting for variation among trainees and supervisors (i.e., potentially unique trainee-or

supervisor-level characteristics that may influence the dependent variables due to the non-independent nature of the observations).

Cross-classified models were planned to first estimate how well the clique predicted trainees' self-reported acquisition of PWCs ("PWC Growth"), as measured by the PTESC growth items. Next, cross-classified models would estimate how well the clique predicted interns' acquisition of PWCs as reported by their supervisors on the TCE ("PWC Competence"). Both sets of models assessed how well supervision competence predicted trainees' acquisition of PWCs while simultaneously accounting for variation among trainees and supervisors. Trainee ID was included as a nesting unit, given that trainees may show individual differences in their response patterns on the PTESC. Supervisor ID was then included as a nesting unit because trainees are nested within supervisors in the dataset.

Four sets of models (Dunn et al., 2015; Hox et al., 2017) were fit for each outcome. The outcome *y* across all equations represents either a trainee-reported PWC Growth score or a supervisor-reported intern PWC Competence score depending on the model in question. The first two equations modeled the variance in the given outcome as a function of variability among (1) trainees, and (2) supervisors. These two models (Equations 1 and 2) use a traditional two-level multilevel modeling approach, whereby observations are nested within only one higher level nesting unit. In Equation 1, the "trainee only" model (ignoring supervisor), the outcome *y* for observation *i* nested by trainee *j* was modeled as:

$$\text{Equation 1: } y_{ij} = \beta_0 + u_{0j} + e_{ij}$$

The fixed parameter β_0 represents the mean outcome score across all trainees, u_{0j} represents the random effect of trainee *j*, and e_{ij} is the residual term for trainee. Results from Equation 1 were used to calculate intraclass correlation coefficients (ICC) for Trainee ID: ICCs represent the

proportion of variance in the outcome due to differences across individuals. The ICC for trainee was generated by dividing the Trainee ID random effect by the total variance.

Similarly in Equation 2, the "supervisor only" model (ignoring trainee), the outcome *y* for observation *i* nested by supervisor *k* was modeled as:

$$\text{Equation 2: } y_{ik} = \beta_0 + u_{0k} + e_{ik}$$

Here, β_0 is the mean outcome score across all supervisors, u_{0k} represents the random effect of supervisor *k*, and e_{ik} is the residual term for supervisor. Results from Equation 2 were used to generate ICCs for Supervisor ID, calculated as the quotient of the Supervisor ID random effect divided by the total variance.

Next, a null cross-classified model (Equation 3) was estimated to examine the components of variance in each nesting unit simultaneously without the predictor. The outcome *y* for observation *i* nested by trainee *j* and supervisor *k* was modeled as:

$$\text{Equation 3: } y_{jk} = \beta_{0jk} + u_{0j} + u_{0k} + e_{jk}$$

In Equation 3, β_{0jk} represents the mean outcome score across all trainees and supervisors, u_{0j} is the random effect for trainee, u_{0k} is the random effect for supervisor, and e_{jk} is the residual for observation *i* nested by trainee *j* and supervisor *k*. The trainee and supervisor ICCs for these null cross-classified models indicate the proportion of variance in the outcome due to Trainee or Supervisor ID, respectively, *while* controlling for the other. Finally, this model was extended in Equation 4 to include the level 2 predictor, supervision competence, which was derived from the EGA:

$$\text{Equation 4: } y_{jk} = \beta_{0jk} + \beta_1 x_{jk} + u_{0j} + u_{0k} + e_{jk}$$

Here, $\beta 1$ represents the fixed effect of supervision competence, and x_{jk} is the supervision competence score for observation *i* nested by trainee *j* and supervisor *k*. Interpretation of the other parameters remains the same. This equation estimates a cross-classified model with supervision competence predicting the given outcome while accounting for nesting of trainees and supervisors. Restricted maximum likelihood (REML) estimation was used across all cross-classified models (Hayes, 2006; Hox et al., 2017). REML is a variant of the maximum likelihood approach. Unlike maximum likelihood estimation, REML produces unbiased estimates of variance and covariance parameters. The following goodness of fit measures were generated: Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and log-likelihood. Values closer to zero indicate better model fit for all indices.

Missing Data

There were 1,301 missing values (5.62% of the data) across all PTESC and TCE responses (203 cases, 114 variables). Missing values included "Not Applicable" responses. Creating seven PTESC domain scales limited the influence of missing values by reducing the number of variables from 114 to 25: there were 157 missing values (3.1% of the data) after creating scales. Listwise deletion (per observation) was used in all of the multilevel models, given that each contained one predictor and one outcome and could not otherwise be estimated.

Results

Descriptive statistics for all variables used in the analyses are reported in Table 2. Means across all variables tended to be negatively skewed, indicating high ratings of supervision competence, trainee growth, and intern competence. The scales created for each of the seven PTESC domains showed strong reliability (Cronbach's alpha values ranging from 0.89-0.98), as shown in Table 3. All domain scales were significantly positively correlated (see Table 4).

Table 2
Descriptive Statistics

Variable	n	Minimum	Maximum	Mean	SD
Domain Scales					
A. Supervisor Competence	203	1.76	5.00	4.69	0.47
B. Diversity	200	1.57	5.00	4.60	0.63
C. Supervisory Relationship	199	1.36	5.00	4.58	0.62
D. Professionalism	198	2.00	5.00	4.71	0.51
E. Evaluation and Feedback	198	1.29	5.00	4.61	0.63
F. Managing Trainee Problems	182	1.33	5.00	4.73	0.62
G. Ethical Conduct	199	1.57	5.00	4.84	0.42
Trainee-Reported DVs					
Research Growth	190	2.00	5.00	4.49	0.68
Ethics Growth	189	3.00	5.00	4.38	0.65
Diversity Growth	190	1.00	5.00	4.35	0.79
Professionalism Growth	190	1.00	5.00	4.64	0.67
Communication Growth	189	1.00	5.00	4.61	0.70
Assessment Growth	189	2.00	5.00	4.60	0.63
Intervention Growth	189	2.00	5.00	4.49	0.76
Supervision Growth	190	1.00	5.00	4.24	0.88
Consultation Growth	190	1.00	5.00	4.46	0.76
Supervisor-Reported DVs					
Research Competence	87	3.00	5.00	4.33	0.44
Ethics Competence	86	3.50	5.00	4.36	0.42
Diversity Competence	86	3.50	5.00	4.41	0.46
Professionalism Competence	87	3.50	5.00	4.47	0.43
Communication Competence	87	3.00	5.00	4.55	0.45
Assessment Competence	84	4.00	5.00	4.40	0.42
Intervention Competence	77	3.63	5.00	4.41	0.41
Supervision Competence	50	3.00	5.00	4.34	0.48
Consultation Competence	77	3.00	5.00	4.37	0.50
Supervision Competence net score	203	0.66	5.00	4.56	0.72

Table 3
Reliability Statistics for PTESC Domain Scales

Domain	*N*	Cronbach's alpha	No. Items
A. Supervisor Competence	94	0.98	29
B. Diversity	178	0.98	14
C. Supervisory Relationship	180	0.95	11
D. Professionalism	180	0.89	7
E. Evaluation and Feedback	185	0.96	14
F. Managing Trainee Problems	76	0.98	6
G. Ethical Conduct	188	0.92	7

Table 4

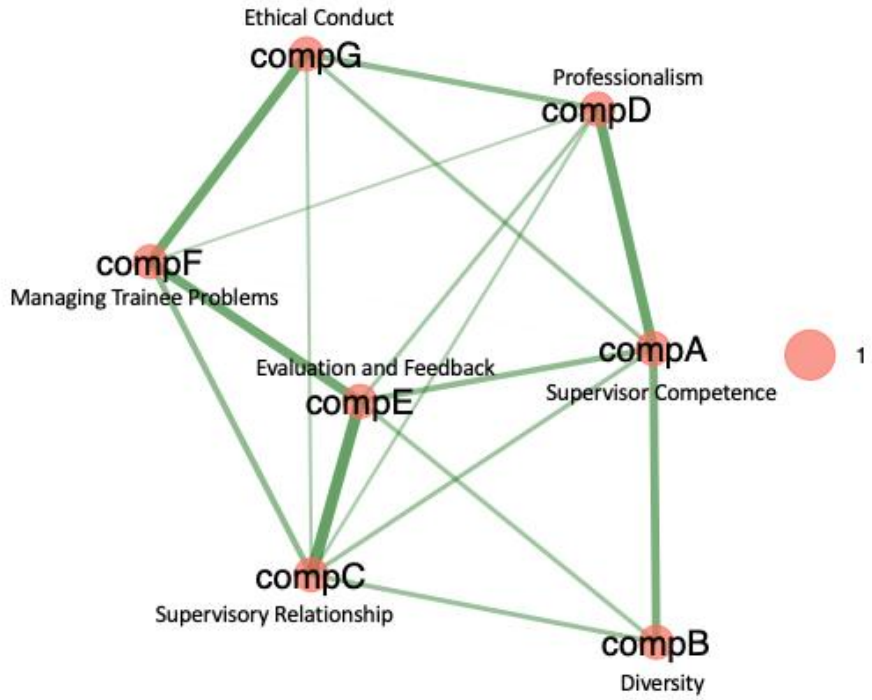
Correlation Coefficients for PTESC Domain Scales and Supervision Competence (SC)

	1	2	3	4	5	6	7
1. Supervisor Competence (A)	—	—	—	—	—	—	—
2. Diversity (B)	0.78**	—	—	—	—	—	—
3. Supervisory Relationship (C)	0.86**	0.75**	—	—	—	—	—
4. Professionalism (D)	0.87**	0.66**	0.82**	—	—	—	—
5. Evaluation and Feedback (E)	0.87**	0.75**	0.92**	0.82**	—	—	—
6. Managing Trainee Problems (F)	0.77**	0.62**	0.83**	0.76**	0.84**	—	—
7. Ethical Conduct (G)	0.78**	0.61**	0.78**	0.78**	0.75**	0.79**	—
8. SC net score	0.67**	0.66**	0.90**	0.84**	0.88**	0.89**	0.79**

*$p < 0.05$; **$p < .01$

Figure 2

PTESC One-Dimension Network



Dimensions of the PTESC

EGA was applied using the PTESC domain scales to answer RQ1: What is the dimensional structure of the PTESC? The EGA revealed a single clique (see Figure 2). This clique/dimension is hereafter labeled ‘Supervision Competence.’ Based on the EGA network plot, the Evaluation and Feedback (Domain E) node appears most central to the Supervision

Competence clique. Heavier edge weights between nodes suggest Evaluation and Feedback is most closely related to Supervisory Relationship (Domain C) and Managing Trainee Problems (Domain F). Managing Trainee Problems is also closely related to Ethical Conduct (Domain G). On the right side of the network plot, Supervisor Competence and Professionalism (Domains A and D) are also connected by a heavier edge weight. RQ2 ("To what extent are the dimensions related to one another?") was not applicable given the one-clique solution achieved by the EGA.

Network loadings (functionally equivalent to factor loadings in factor analysis) are presented in Table 5. The domains with the highest network loadings were Evaluation and Feedback (0.448), Supervisor Competence (0.428), and Supervisory Relationship (0.428). Generally, the network loadings were low to moderate; however, sample size must be considered when interpreting factor loadings. For a sample of 300, Tabachnick and Fidell (2007) suggest minimal loadings of at least 0.32 or greater for meaningful interpretation. The current sample size of 203 is smaller, yet five of the seven domain scales meet this threshold; only Diversity and Ethical Conduct (Domains B and G) fall below this cutoff with loadings of 0.22 and 0.30, respectively.

Results of the bootstrapping with 500 replications indicated that the EGA produced a one-dimension model 86.6% of the time, a two-dimension model 12.4% of the time, and a three-dimension model 1.0% of the time. Table 6 shows the percentage of the time that each domain scale was assigned to the competing dimensional models over the course of 500 replications; each was assigned to the single dimension of Supervision Competence 91-98% of the time. The Supervisory Relationship (Domain C) was the most consistent in its assignment to the one-dimension solution (98% of the time). Professionalism (Domain D) was the least consistent in its assignment: it was assigned to the one-dimension solution 91% of the time, and it was assigned

to a second dimension in a two-dimension solution 9.2% of the time.

Table 5

Network Loadings onto the Supervision Competence Clique

	1
A. Supervisor Competence	0.428
B. Diversity	0.222
C. Supervisory Relationship	0.428
D. Professionalism	0.337
E. Evaluation and Feedback	0.448
F. Managing Trainee Problems	0.361
G. Ethical Conduct	0.301

Table 6

Item-Dimension Assignment with 500 Replications

	1	2	3
A. Supervisor Competence	0.952	0.048	
B. Diversity	0.952	0.048	
C. Supervisory Relationship	0.976	0.022	0.002
D. Professionalism	0.908	0.092	
E. Evaluation and Feedback	0.970	0.028	0.002
F. Managing Trainee Problems	0.918	0.074	0.008
G. Ethical Conduct	0.928	0.064	0.008

Table 7

CFA Results

Latent Variable:	Estimate	Std.Err.	z-value	*p*	Std.all
Supervision Competence					
A. Supervisor Competence	1.000				0.947
B. Diversity	1.126	0.101	11.197	0.000	0.778
C. Supervisory Relationship	1.361	0.100	13.577	0.000	0.952
D. Professionalism	1.040	0.067	15.447	0.000	0.886
E. Evaluation and Feedback	1.376	0.091	15.188	0.000	0.949
F. Managing Trainee Problems	1.200	0.101	11.896	0.000	0.859
G. Ethical Conduct	0.819	0.095	8.582	0.000	0.842
Variances:					
A. Supervisor Competence	0.023	0.006	4.077	0.000	0.103
B. Diversity	0.165	0.038	4.302	0.000	0.394
C. Supervisory Relationship	0.039	0.009	4.503	0.000	0.094
D. Professionalism	0.060	0.009	6.621	0.000	0.216
E. Evaluation and Feedback	0.042	0.010	4.071	0.000	0.099
F. Managing Trainee Problems	0.103	0.025	4.187	0.000	0.263
G. Ethical Conduct	0.055	0.015	3.607	0.000	0.291
Supervision Competence	0.201	0.060	3.355	0.001	1.000

Results of the CFA are presented in Table 7. Model fit indices from the CFA analysis indicated good fit for the one-factor model: $\chi^2 = 0.395$ (14), CFI = 1.00, RMSEA = 0.00 with 90% CI [0.00, 0.00], SRMR = 0.027. Standardized factor loadings for each scale were strong (range of 0.78-0.95) and statistically significant at $p < .001$.

Net scores (functionally equivalent to factor scores in factor analysis) were generated for the one-clique model to be used as the predictor in the subsequent cross-classified multilevel models. These scores are referred to as Supervision Competence (SC) net score. Correlations between SC net score and (a) the PTESC domain scales, and (b) the outcome variables used in the cross-classified multilevel models are reported in Tables 4 and 8, respectively. Differences in the nature and raters of the trainee- and supervisor-reported outcomes are reflected in these variables' correlation coefficients: all trainee-reported Growth outcomes showed significant positive correlations with one another (range of 0.18-0.79) and with SC net score (range of 0.39-0.68). All supervisor-reported Competence outcomes showed significant positive correlations with one another (range of 0.54-0.81), but none were significantly correlated with SC net score. Between the nine Growth and nine Competence outcomes, there were only five significant correlations: Growth in Intervention was negatively correlated with Ethics Competence (-0.28); Growth in Professionalism was negatively correlated with Competence in Assessment (-0.25), Intervention (-0.25), and Supervision (-0.32); and Growth in Consultation was positively correlated with Consultation Competence (0.25).

Multilevel Modeling Results

Cross-classified multilevel modeling (CCMLM) was used to address RQ3: How well do(es) the dimension(s) of supervision competence predict trainee acquisition of the nine profession-wide competencies (PWCs)? CCMLM models assessed how well SC net score

predicted trainee-reported Growth outcomes and supervisor-reported Competence outcomes while simultaneously accounting for variation among trainees and supervisors. Trainee ID and Supervisor ID were both included as nesting units. Results from Equation 1 were used to calculate intraclass correlation coefficients (ICC; the proportion of variance in the outcome due to differences across individuals) for Trainee ID: ICCs for trainee ranged from 0.02-0.38 in null models for trainee-reported Growth outcomes and from < 0.00-0.18 in null models for supervisor-reported Competence outcomes. Results from Equation 2 were used to generate ICCs for Supervisor ID: ICCs for supervisor ranged from < 0.00-0.22 in null models for trainee-reported Growth outcomes and from 0.28-0.58 in null models for supervisor-reported Competence outcomes. The trainee and supervisor ICCs for the Equation 3 null cross-classified models indicate the proportion of variance in the outcome due to trainee or supervisor ID, respectively, *while* controlling for the other. All ICCs are reported in Table 9.

Predicting PWC Growth as Self-Reported by Trainees

Results of the CCMLM procedure for models predicting trainee-reported PWC Growth outcomes are reported in Table 10. Sufficient variation across trainees and supervisors (as indicated by ICCs) warranted controlling for trainee and supervisor in the cross-classified models predicting Growth in Diversity, Intervention, Supervision, and Consultation; Equations 1-4 were estimated as specified in the Methods for these four outcomes. In each of these models, Supervision Competence was a significant positive predictor of PWC Growth while accounting for both trainee and supervisor identity.

When predicting Growth in Research, Ethics, Professionalism, and Assessment, the ICCs for Supervisor ID generated by Equation 2 indicated that the variance in the outcome due to the supervisor was near 0. Due to lack of variation across supervisors, supervisor identity was

removed from these cross-classified models (Equation 4) to simplify the models and improve fit. The results indicated that Supervision Competence was a significant positive predictor of Growth in each of these PWCs while accounting for trainee identity.

Finally, ratings for Growth in Communication showed little variation across trainees or supervisors (ICCs for Trainee ID and Supervisor ID were near 0), indicating a multilevel modeling approach accounting for the nesting of these variables was not necessary. Therefore, a simple linear regression model was used to predict this outcome. The model was statistically significant ($R^2 = 0.363$, $F(1, 187) = 106.68$, $p < 0.001$). In summary, supervision competence was a significant positive predictor of growth across all nine PWCs as reported by trainees.

Predicting Intern Competence as Reported by Supervisors

Results of the CCMLM procedure for models predicting supervisor-reported intern Competence outcomes are reported in Table 11. Equations 1-4 were estimated as specified in the Methods for four of the nine outcomes: Competence in Research, Diversity, Professionalism, and Intervention. Supervision Competence did not significantly predict any of these outcomes.

When predicting interns' Competence in Ethics, Communication, Assessment, Supervision, and Consultation, the ICC for Trainee ID generated by Equation 1 indicated that the variance in the outcome due to the trainee was near 0. Therefore, trainee identity was removed from these cross-classified models (Equation 4) to simplify them and improve fit. Only the model predicting Ethics Competence was significant: Supervision Competence predicted lower scores of interns' Ethics Competence.

Table 8

Correlation Coefficients for Outcome Variables and Supervision Competence (SC)

	Growth in…									Competence in…								
	1	2	3	4	5	6	7	8	9	1	2	3	4	5	6	7	8	9
Growth in…																		
1	—																	
2	0.51**	—																
3	0.39**	0.45**	—															
4	0.48**	0.48**	0.54**	—														
5	0.44**	0.44**	0.47**	0.79**	—													
6	0.52**	0.41**	0.36**	0.53**	0.45**	—												
7	0.21**	0.22**	0.34**	0.45**	0.40**	0.18*	—											
8	0.48**	0.45**	0.46**	0.53**	0.49**	0.46**	0.31**	—										
9	0.38**	0.40**	0.41**	0.53**	0.46**	0.45**	0.45**	0.49**	—									
Competence in…																		
1	-0.10	0.04	0.01	-0.06	-0.13	-0.08	-0.09	0.05	0.14	—								
2	0.02	0.05	-0.04	-0.19	-0.20	0.03	-0.28*	-0.06	0.04	0.65**	—							
3	0.11	0.13	-0.01	-0.13	-0.11	0.05	-0.15	0.19	0.16	0.59**	0.69**	—						
4	0.09	0.09	-0.01	-0.08	-0.03	0.04	-0.16	0.09	0.11	0.73**	0.74**	0.66**	—					
5	0.05	0.16	-0.01	-0.01	-0.02	0.04	-0.14	0.18	0.07	0.56**	0.63**	0.65**	0.72**	—				
6	-0.11	0.02	-0.01	-0.25*	-0.22	-0.06	-0.15	-0.11	0.06	0.68**	0.69**	0.65**	0.68**	0.64**	—			
7	-0.01	0.07	0.07	-0.25*	-0.16	0.02	-0.15	0.01	0.03	0.65**	0.76**	0.72**	0.76**	0.75**	0.81**	—		
8	0.13	0.11	0.07	-0.32*	-0.22	0.15	-0.11	-0.22	0.25	0.66**	0.66**	0.70**	0.55**	0.54**	0.66**	0.67**	—	
9	0.11	0.08	0.02	-0.07	-0.08	0.11	-0.19	0.16	0.25*	0.63**	0.60**	0.66**	0.59**	0.62**	0.56**	0.74**	0.70**	—
SC	0.44**	0.51**	0.53**	0.68**	0.60**	0.39**	0.44**	0.56**	0.45**	-0.05	-0.20	0.04	-0.07	-0.05	-0.12	-0.02	-0.13	0.11

Note. 1. Research 2. Ethics 3. Diversity 4. Professionalism 5. Communication 6. Assessment 7. Intervention 8. Supervision 9. Consultation

*$p < 0.05$, **$p < .01$

Table 9

Null Models and Intraclass Correlation Coefficients (ICCs)

		Growth Outcome Models		Competence Outcome Models	
Outcome	Null Model	Trainee ICC	Supervisor ICC	Trainee ICC	Supervisor ICC
Research	1. Trainee Only	0.37	—	0.06	—
	2. Supervisor Only	—	0.00	—	0.44
	3. Cross-Classified	0.37	0.00	0.14	0.49
Ethics	1. Trainee Only	0.30	—	0.00	—
	2. Supervisor Only	—	0.00	—	0.48
	3. Cross-Classified	0.30	0.00	0.04	0.48
Diversity	1. Trainee Only	0.38	—	0.05	—
	2. Supervisor Only	—	0.08	—	0.46
	3. Cross-Classified	0.48	0.20	0.38	0.31
Professionalism	1. Trainee Only	0.05	—	0.18	—
	2. Supervisor Only	—	0.06	—	0.28
	3. Cross-Classified	0.11	0.09	0.27	0.30
Communication	1. Trainee Only	0.02	—	0.00	—
	2. Supervisor Only	—	0.07	—	0.43
	3. Cross-Classified	0.05	0.08	0.05	0.43
Assessment	1. Trainee Only	0.17	—	0.00	—
	2. Supervisor Only	—	0.00	—	0.51
	3. Cross-Classified	0.19	0.02	0.07	0.51
Intervention	1. Trainee Only	0.27	—	0.02	—
	2. Supervisor Only	—	0.22	—	0.58
	3. Cross-Classified	0.12	0.21	0.12	0.58
Supervision	1. Trainee Only	0.24	—	0.00	—
	2. Supervisor Only	—	0.05	—	0.51
	3. Cross-Classified	0.28	0.06	0.04	0.50
Consultation	1. Trainee Only	0.16	—	0.00	—
	2. Supervisor Only	—	0.04	—	0.44
	3. Cross-Classified	0.19	0.06	0.07	0.44

Table 10

Coefficients for Cross-classified Models Predicting Trainee-Reported Growth in Profession-Wide Competencies

Growth Outcome	n	Intercept (s.e.)	Supervision Competence (s.e.)	Trainee Random Effect	Supervisor Random Effect	Residual s.d.	AIC	BIC	Log Lik.
Research	obs = 190 groups (tr) = 101	1.93*** (0.40)	0.56*** (0.09)	0.35		0.50	351.4	364.4	-171.70
Ethics	obs = 189 groups (tr) = 100	1.65*** (0.37)	0.59*** (0.08)	0.24		0.51	325.3	338.2	-158.64
Diversity	obs = 190 groups (tr) = 101 groups (sup) = 31	0.64 (0.45)	0.81*** (0.09)	0.41	0.44	0.47	387.8	404.0	-188.91
Professionalism	obs = 190 groups (tr) = 101	0.51 (0.32)	0.90*** (0.07)	0.18		0.46	283.4	296.4	-137.69
Communication [a]	obs = 189	0.91* (0.36)	0.80*** (0.08)				319.0	328.7	-156.48
Assessment	obs = 189 groups (tr) = 101	2.36*** (0.39)	0.48*** (0.08)	0.13	0.17	0.55	346.3	362.5	-168.15
Intervention	obs = 189 groups (tr) = 100 groups (sup) = 31	1.84*** (0.43)	0.58*** (0.09)	0.24	0.30	0.55	379.1	395.3	-184.55
Supervision	obs = 190 groups (tr) = 101 groups (sup) = 31	-0.40 (0.48)	0.99*** (0.10)	0.20	0.39	0.59	420.0	436.3	-205.03
Consultation	obs = 190 groups (tr) = 101 groups (sup) = 31	1.30** (0.45)	0.68*** (0.10)	0.19	0.29	0.58	398.1	414.3	-194.03

Note. Tr = trainee. Sup = supervisor.

[a] The model predicting Growth in Communication used simple linear regression.

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Supplemental Analyses

Models Including Demographic Variables

Additional analyses were conducted to examine the potential influence of trainee demographics in the 10 models that produced significant results (i.e., the nine PWC Growth models and the intern Ethics Competence model). The demographic variables examined were gender, race/ethnicity, and trainee position, each of which was coded categorically. Gender was coded as female or male. Race/ethnicity categories were condensed to either White or person of color due to the small cell sizes for specific non-White racial/ethnic groups. Trainee position was coded as extern, intern, or postdoctoral fellow. Each of the 10 models was re-estimated three separate times to examine the fixed effect of each demographic variable individually.

Gender had no significant effect in any of the models. Race/ethnicity had a significant effect only in the model predicting Diversity Growth: White racial/ethnic identity predicted higher scores on Diversity Growth. Supervision Competence remained a significant positive predictor in this model. Model statistics and fit indices are reported in Table 12: the model including race/ethnicity showed better fit than the original model. Trainee position had a significant effect only in the model predicting Intervention Growth: being a postdoctoral fellow predicted lower scores on Intervention Growth. Supervision Competence remained a significant positive predictor in this model, as well. Including trainee position in this model had no clear impact (i.e., improving or worsening) on model fit (see Table 12).

Exploring Other Outcomes: Supervisor-Reported Change in Competence

Although interns' baseline competency levels (i.e., at the start of the training year) were not assessed, supervisors did evaluate interns' PWC acquisition using the TCE at midyear (6 months). Midyear mean ratings (Competence scores) for each PWC domain were available for

67 of the 87 intern-supervisor dyads included in the previous analyses. For each of the nine PWCs, the midyear mean ratings were subtracted from the endpoint mean ratings to generate nine new variables representing Change in Competence from the 6- to 12-month mark in the training year. These variables were created as an alternate supervisor-reported outcome more closely aligning conceptually with the trainee-reported PWC Growth variables.

The CCMLM procedures were repeated using these nine variables as the outcome. The models predicting Change in Competence in Research, Assessment, Intervention, Supervision, and Consultation were not significant. When predicting Change in Competence in Ethics and Diversity, the ICCs for Trainee ID generated by Equation 1 were near 0, so trainee identity was removed from the cross-classified models (Equation 4) to improve parsimony and fit. The cross-classified model predicting Change in Ethics approached significance ($p = 0.066$) with Supervision Competence trending toward a positive predictor. The cross-classified model predicting Change in Diversity was nonsignificant. Simple regression models (not accounting for supervisor identity) were subsequently tested for model simplification purposes, given the small sample size. The simple regression models predicting Change in Ethics ($R^2 = 0.053$, $F(1, 64) = 3.596$, $p = 0.062$) and Change in Diversity ($R^2 = 0.055$, $F(1, 65) = 3.798$, $p = 0.056$) continued to approach significance.

When predicting Change in Competence in Professionalism and Communication, the ICCs for Supervisor ID generated by Equation 2 were near 0, so supervisor identity was removed from the cross-classified models (Equation 4). The cross-classified model predicting Change in Professionalism approached significance ($p = 0.052$) with Supervision Competence trending toward a positive predictor. The cross-classified model predicting Change in Communication was nonsignificant. Again, simple regression models (not accounting for trainee identity) were

Table 11

Coefficients for Cross-classified Models Predicting Supervisor-Reported Intern Competence

Competence Outcome	n	Intercept (s.e.)	Supervision Competence (s.e.)	Trainee Random Effect	Supervisor Random Effect	Residual s.d.	AIC	BIC	Log Lik.
Research	obs = 87 groups (tr) = 23 groups (sup) = 21	4.38*** (.39)	-0.02 (0.08)	0.17	0.32	0.30	101.6	113.9	-45.81
Ethics	obs = 86 groups (sup) =21	5.24*** (0.33)	-0.19** (0.07)		0.32	0.30	83.0	92.8	-37.49
Diversity	obs = 86 groups (tr) = 23 groups (sup) =21	3.59*** (0.67)	0.17 (0.14)	0.15	0.33	0.31	102.0	114.2	-45.98
Professionalism	obs = 87 groups (tr) = 23 groups (sup) =21	5.00*** (0.37)	-0.12 (0.08)	0.24	0.24	0.28	92.6	104.9	-41.30
Communication	obs = 87 groups (sup) =21	4.59*** (0.39)	-0.02 (0.08)		0.32	0.35	107.7	117.5	-49.84
Assessment	obs = 84 groups (sup) =20	4.65*** (0.56)	-0.05 (0.12)		0.32	0.30	82.0	91.7	-37.00
Intervention	obs = 77 groups (tr) = 23 groups (sup) =20	3.79*** (0.57)	0.14 (0.12)	0.15	0.33	0.22	60.7	72.4	-25.38
Supervision	obs = 50 groups (sup) =16	4.63*** (0.81)	-0.07 (0.17)		0.35	0.34	67.0	74.7	-29.50
Consultation	obs = 77 groups (sup) =20	3.14*** (0.71)	0.25+ (0.15)		0.34	0.36	100.5	109.9	-46.24

Note. Tr = trainee. Sup = supervisor.

+$p < 0.10$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

Table 12

Coefficients for Cross-classified Models including Demographic Variables

Growth Outcome	n	Intercept (s.e.)	Supervision Competence (s.e.)	Trainee Random Effect	Supervisor Random Effect	**Race/Ethn.** (White)	Residual s.d.	AIC	BIC	Log Lik.
Diversity										
Model 1	obs = 190 groups (tr) = 101 groups (sup) = 31	0.64 (0.45)	0.81*** (0.09)	0.41	0.44	—	0.47	387.8	404.0	-188.91
Model 2	obs = 169 groups (tr) = 84 groups (sup) =30	0.17 (0.60)	0.85*** (0.13)	0.18	0.15	0.28* (0.14)	0.45	340.7	359.5	-164.34
						Position (Fellow)				
Intervention										
Model 1	obs = 189 groups (tr) = 100 groups (sup) = 31	1.84*** (0.43)	0.58*** (0.09)	0.24	0.30	—	0.55	379.1	395.3	-184.55
Model 2	obs = 189 groups (tr) = 100 groups (sup) = 31	1.95*** (0.44)	0.57*** (0.09)	0.05	0.10	-0.68* (0.32)	0.55	381.4	404.0	-183.68

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$

tested. In the simple regression models, Supervision Competence was a significant positive predictor of Change in Professionalism ($R^2 = 0.062$, $F(1, 65) = 4.292$, $p < 0.05$) and Change in Communication ($R^2 = 0.080$, $F(1, 65) = 5.622$, $p < 0.05$).

Discussion

This study applied EGA to analyze the dimensional structure of the PTESC, a trainee-report measure of supervision competence, and used the resulting single dimension to predict trainee acquisition of PWCs. The findings showed supervision competence predicted trainee-reported growth across all nine PWCs and positive change in professionalism and communication competence as reported by supervisors. Supervision competence predicted lower ratings of interns' competence in ethics. Race/ethnicity impacted trainees' reports of their growth in diversity competence. The study's strengths and limitations and implications for research, theory, and practice are further explored.

Major Findings

One Dimension of Supervision Competence

The EGA revealed a single dimension of Supervision Competence. Based on the network plot and network loadings, the domains most strongly related to the Supervision Competence clique were: Evaluation and Feedback, Supervisory Relationship, and Supervisor Competence. Diversity and Ethical Conduct were the two weakest domains; however, all seven domains were reliably assigned to the single dimension of Supervision Competence more than 90% of the time and had strong positive standardized factor loadings in the CFA. The results indicate that all seven domains meaningfully contribute to supervision competence and therefore should be attended to in training and evaluation of supervision.

The Evaluation and Feedback domain was most strongly related to Supervision Competence, which makes sense given that supervision is an inherently evaluative relationship. The Supervisory Relationship's strong association with Supervision Competence was also unsurprising, as the relationship has been called "the very heart and soul of supervision" (Watkins, 2014, p. 20). The Supervisor Competence domain on the PTESC is unique in that it includes two goal items: one focused on competence in the clinical services being provided, and one focused on competence in provision of supervision (e.g., seeking supervision competence through education and training). The name of this domain is somewhat misleading, as the APA (2014) *Guidelines* and this study show that this one domain alone does not comprise all aspects of a supervisor's competence. Nonetheless, this domain's strong relationship to the Supervision Competence clique speaks to the importance of a supervisor's clinical competence, as well as their training and valued engagement in the practice of supervision (Ladany et al., 2013; Magnuson et al., 2000).

The relatively lower loadings of Diversity and Ethical Conduct align with their visual representation in the EGA network plot: these two nodes fall at the outermost points of the plot, whereas the other five domain nodes appear more central. It is possible that these domains represent broader, more foundational, values-based competencies that have some conceptual differences from more functional, skill-based competencies. Foundational competencies can be thought of as the knowledge, skills, attitudes, or values that provide the foundation for a psychologist's professional activities, whereas functional competencies involve the major functions a psychologist performs (e.g., assessment, psychotherapy, consultation; Fouad et al., 2009). The Competencies Conference supervision work group (Falender et al., 2004) included valuing of ethics and sensitivity to diversity in the core "values" competencies of supervision.

Appreciation for diversity and ethical, values-based practice are also two superordinate values in the competency-based supervision framework put forth in this work (Figure 1). Additionally, the PTESC items in these two domains emphasize values and attitudes (e.g., "values and models ethical behavior," "recognizes the value of and pursues ongoing training in diversity"), as opposed to more skills-based competencies like assessment, intervention, and consultation.

The Diversity domain's representation in the EGA network is also interesting in the context of broader discussions in the field about diversity/multicultural competence: some advocate for a focus on a multicultural *orientation* or lens (e.g., Watkins et al., 2019). Diversity or multicultural competence models have typically emphasized the requisite knowledge and skills one must possess to work effectively across cultures and identities, but Watkins and colleagues (2019) posit that attitudes and values are at the core of multicultural competence and have received insufficient attention. They propose a multicultural orientation framework to address this gap comprised of three components: cultural humility, cultural comfort, and cultural opportunities. The cultural humility and comfort components, in particular, represent attitudes rather than knowledge or skills. The multicultural orientation framework is "grounded in the core conviction that culture matters in society" (Watkins et al., 2019, p. 40), which is a values-based belief. Unique considerations about the nature of this area of competence may be reflected in the Diversity domain's relative distance from the other nodes and lower network loading.

Like Diversity, Ethical Conduct was the second node that was more weakly related to Supervision Competence, and it may be thought of as a foundational, values-based domain. This conceptual difference may have influenced its representation in the EGA network. It is also possible that ethical conduct is less central to trainees' perceptions of what makes a supervisor competent; perhaps it is taken for granted that a supervisor should behave ethically. It is also of

note that the Ethical Conduct domain scale had the highest mean (m = 4.84 on a scale of 5.0) and the lowest standard deviation (SD = 0.42) of all the scales. The little variation in scores may have affected its representation in the network and/or its relation to the Supervision Competence clique. However, the EGA bootstrapping and the WLSMV estimator used in the CFA addressed issues of non-normality in data (Finney & DiStefano, 2006).

Although the network loadings of the Diversity and Ethical Conduct domains were weak, these domains were nonetheless very consistent in their assignment to the single dimension solution, and all seven domains showed a strong positive correlation with Supervision Competence. Furthermore, model fit indices from the CFA indicated good fit for the one-factor model, and standardized factor loadings for each domain scale were strong and statistically significant at $p < .001$. These findings lend empirical support to (a) the APA's (2014) seven domains of supervision competence, and (b) the PTESC as a measure of these competency areas. They also highlight the importance of attending to all domains of supervision competence in training and evaluation—attending to the supervisory relationship alone is necessary but insufficient for provision of competent supervision.

Does Supervision Competence Predict Trainee Competence?

Supervision Competence predicted trainee-reported growth in all nine PWCs: Research, Ethics, Diversity, Professionalism, Communication, Assessment, Intervention, Supervision, and Consultation. This is the first known study providing evidence that supervision competence predicts growth in trainees' professional competencies. These results support previous literature indicating supervision positively influences trainee skill acquisition (Bambling, 2014; Wheeler & Richards, 2007) and make a significant contribution in highlighting the wide range of trainee competencies influenced by competent supervision. Further, trainees' self-assessments of their

own competency growth should not be dismissed: recent research has shown no significant differences between trainee and supervisor ratings of trainee competencies (Hitzeman et al., 2019). This study's findings emphasize how powerful supervision can be as a teaching method; in growing trainees' competencies, the competent supervisor not only prepares the trainee for eventual independent practice but protects the public.

The only significant finding from the models predicting supervisor-reported competence outcomes was that Supervision Competence negatively predicted interns' Competence in Ethics. Ethics Competence ratings varied little across trainees, so this model only accounted for supervisor identity. The ICC for supervisor identity across the null models indicated 48% of variance in Ethics Competence was due to the supervisor. Because the supervisor was also the rater, this value suggests notable subjectivity by supervisors in Ethics ratings. The inverse relationship between Supervision Competence and interns' Ethics Competence might initially suggest that trainees of more competent supervisors engage in less ethical behavior. However, it seems more likely that competent supervisors are more stringent—and maybe appropriately so—in their ratings of trainees' Ethics Competence. It has been well-documented that supervisor ratings of trainee competence are subject to halo and leniency effects (e.g., Bogo et al., 2002; Gonsalvez & Crowe, 2014; Gonsalvez & Freestone, 2007; Gonsalvez et al., 2021; Strom et al., 2016). More competent supervisors' lower ratings of trainee Ethics Competence may reflect more effective gatekeeping, an essential function of supervision.

Differences in Trainee- versus Supervisor-Reported Outcomes. Although they are both related to the PWCs, the trainee-reported Growth variables and supervisor-reported Competence variables are quite different in nature. The trainee-reported growth items from the PTESC state: "My training experience with this supervisor facilitated the acquisition of

competency in [PWC].” These variables assess trainee-reported effectiveness of supervision in terms of its impact on their own competencies. The supervisor-reported Competence variables, on the other hand, represent the mean of the supervisor’s ratings for the intern on each PWC domain of the TCE. These variables measure the intern’s level of competence at a particular point in time without accounting for their baseline or previous level of competence. These differences may explain why Supervision Competence predicted growth across all PWCs but yielded minimal findings in terms of interns’ competence levels. Each trainee begins the supervisory relationship with varying strengths and weaknesses in their domains of competence based on individual differences, lived experiences, and variations in prior training. Differences in baseline competence levels, which likely influence later competence levels, are not accounted for by TCE ratings; however, the PTESC growth items attempt to capture the trainee’s pre- and post-changes in competence within a cross-sectional design.

The question of whether supervision competence predicts trainee acquisition of PWCs might be better assessed longitudinally, where one can compare trainee competence before and after working with a particular supervisor. However, even in a longitudinal study, the PTESC growth items would remain valuable. They provide a unique, face valid impression of the impact of a particular supervisor’s competence, which is difficult to parse when the trainee is receiving supervision from multiple supervisors. Additionally, trainee perceptions may be particularly useful in measuring progress (Larkin & Morris, 2015). Linear regression models predicting change in Ethics and Diversity competence were also trending toward significance.

The supplemental analyses using Change in Competence outcomes aimed to test the hypothesis that the models predicting supervisor-reported outcomes were mostly null because they did not account for the intern’s baseline competence levels. Simple linear regression models

indicated Supervision Competence predicted positive changes in interns' competence in Professionalism and Communication. Models predicting change in Ethics and Diversity were trending toward significance. Interestingly, these findings suggest supervisors perceive that supervision competence has more influence on trainees' foundational competencies than functional competencies. In future research, it is worth exploring and comparing the influence of supervision competence on foundational versus functional competencies from both the trainee and supervisor perspective.

Finally, the null results of the models predicting supervisor-reported Competence outcomes (all except Ethics) should also be considered in light of the limited sample size. Competence outcomes were only available for the intern subset of the sample: 24 interns were assessed by 21 supervisors to produce 87 matched TCE cases for analysis. It is worth re-running these analyses after additional data collection to assess whether a more robust sample would yield different results. In sum, this study's findings suggest that supervision competence is a better predictor of trainees' growth or change in competence than current level of competence. This makes sense, given each trainee has their own unique strengths and weaknesses. However, further research with a larger sample is warranted to clarify this conclusion and to assess for differential influence of supervision competence on particular PWCs.

The Influence of Trainee Demographics

Trainee demographic variables were added to the cross-classified models with significant findings to explore their potential influence, in conjunction with Supervision Competence, on trainee competence growth. White racial/ethnic identity predicted greater growth in Diversity. Although White trainees may possess other marginalized identity characteristics, it seems plausible that trainees not belonging to a historically excluded racial/ethnic group might have

more “room to grow” in terms of diversity/multicultural competence. In a mixed methods study of 397 clinical psychology doctoral students, Gregus et al. (2019) found that only White students reported that their training programs provided safe and responsive environments. White trainees’ Diversity growth ratings, therefore, may also reflect their perceptions of a safe environment to have culturally relevant discussions.

It also makes sense that trainees of color might provide more modest ratings when asked about attributing growth in Diversity competence to their work with a supervisor, as research has shown racial/ethnic differences in trainees’ perceptions of multicultural training. Gregus and colleagues (2019) found that Black clinical psychology doctoral students perceived significantly less support for multicultural training in supervision compared to their peers of other races, and they perceived that faculty were less supportive of multicultural discussions. These differences showed large effect sizes. Interestingly, Black students’ perceptions were significantly lower than those of Asian and multiracial students, suggesting important differences across racial/ethnic groups worth exploring with a larger sample. Relatedly, using qualitative analysis to explore racial dynamics in supervision dyads of color, Jernigan et al. (2010) identified racial identity as being just as important as race in its influence on perceptions of diversity-related discussions in supervision. Finally, it is an unfortunate truth that trainees of color may experience racial/ethnic microaggressions in supervision (Constantine & Sue, 2007; Murphy-Shigematsu, 2010; Sue et al., 2007). Such experiences might also contribute to more modest ratings of Diversity growth from trainees of color compared to their White peers.

In looking at trainee position/training level as a demographic variable, being a postdoctoral fellow negatively predicted growth in Intervention. There were only three postdocs in this sample, so this finding should be interpreted with extreme caution. However, if replicated

with a larger sample, this finding would support previous research suggesting that competency development among trainees and psychologists plateaus over time (e.g., Price et al., 2017; Tracey et al., 2014). Research has shown that more advanced trainees report greater competence in intervention skills than those earlier in training (Kamen et al., 2010); however, intervention competence may plateau around year 3 of graduate school (Larkin & Morris, 2015). Postdocs near the end of their training and on the precipice of independent practice may perceive supervision as contributing less to their growth as clinicians because they have more advanced intervention skills, and therefore less to learn compared to more novice trainees. Postdocs might also perceive supervision as contributing less to their growth in intervention skills simply because they receive less of it: the Standards of Accreditation require interns to receive twice as many supervision hours as postdocs (APA, 2015). Finally, trainee gender had no effect across models. This variable was limited by its binary coding; future research would benefit from exploring the influence of trainee gender in perceptions of supervision across the spectrum of gender identities.

Strengths

Despite the "culture of competence" (Roberts et al., 2005, p. 356) adopted by health service psychology, empirical support has lagged behind theoretical advancements and this shift in the approach to training (Callahan & Watkins, 2018). This study makes several significant contributions to the field. First, the findings show that supervision competence is a measurable construct and demonstrate the utility of the PTESC for assessing supervision competence from the trainee perspective. This is the first known study to lend empirical support to the seven domains of the APA's (2014) *Guidelines for Clinical Supervision*, indicating that all seven domains contribute to supervision competence.

This is also the first known study to demonstrate the influence of supervision competence on trainees' competency development. Furthermore, findings were strengthened and enriched by use of outcomes from multiple reporters. Supervision competence predicted trainee self-reported growth in all nine PWCs and supervisor-reported growth in trainees' professionalism and communication competencies. Supervision competence predicted lower ratings of interns' ethics competence, which may reflect more effective gatekeeping of competent supervisors. These findings collectively highlight why the development and evaluation of supervision competence are so important to health service psychology as a whole.

The robust statistical methods used also represent a strength of this study. EGA with bootstrapping was used to derive the one-dimension solution of the PTESC. EGA is a recently developed psychometric network technique, which is highly accurate in estimating the dimensional structure of a dataset and performs as well as, if not better than, traditional types of factor analysis (Christensen et al., 2019; Golino et al., 2020). Cross-classified multilevel modeling accounted for the nested nature of the dataset and the fact that trainees worked with multiple supervisors; trainee and supervisor identities were controlled for whenever appropriate in predicting outcomes.

Limitations

The present study is limited somewhat by the sample size. Due to the number of observations (n = 203) and the number of items on the PTESC (n = 96, not including growth items), the planned item-level analysis using multilevel exploratory factor analysis (EFA) could not be conducted. Use of EFA would have allowed for item reduction to refine the measure and improve its utility (i.e., reduce rater burden). However, EGA was an appropriate, and likely even better (Golino & Epskamp, 2017), alternative analysis for examining the dimensional structure of

the measure. The sample size primarily limited the conclusions that could be drawn from analyses involving TCE outcomes (intern Competence scores, n = 87 matched dyads; intern Change in Competence scores, n = 67 matched dyads). In particular, the Change in Competence outcomes trending toward significance warrant re-evaluation after additional data collection.

The PTESC is a measure of supervision competence from the trainee perspective. Subjective perceptions involve inherent bias, and in this case, may have been impacted by impression management. The power differential in the supervisory relationship is powerful and must be acknowledged. The UCM Training Director informed externs and postdocs that their responses were optional, anonymous, and would only be shared with supervisors in aggregate form. However, they still may have feared being identified, particularly when evaluating a supervisor who has few trainees. Interns' evaluations were required and submitted to the Training Director with their names, which were subsequently removed. This lack of initial anonymity could have heightened impression management and/or fears of retaliation for providing negative feedback.

The skewness of the PTESC domains suggests trainees generally perceived their supervisors to be highly competent. It may also reflect, in part, this issue of the power dynamics in supervision. The skew is also likely influenced by bias in the Likert-scale ratings. Trainees rated the frequency with which the supervisor displayed particular behaviors using the following response options: *Never, Rarely, Frequently, Typically,* or *Almost Always*. The central option of *Frequently* is non-neutral and therefore pulls for negative skew in the data. The difference between observing a behavior *Frequently, Typically,* and *Almost Always* may also be difficult to discern and report. The influence of skew in EGA is unknown; however, the bootstrapping procedure, as well as the WLSMV estimation method used in the CFA, are robust to deviations

from normality in the dataset. Nonetheless, future research using the PTESC should make all efforts to support trainees' honest and/or anonymous reporting and account for the biased response options.

Psychometric properties of the TCE are unknown, and their exploration is beyond the scope of this study. Until recently, no validated measure of trainee competencies existed. Price and colleagues (2017) used item response theory to revise and validate the Practicum Evaluation Form (PEF; University of North Texas Psychology Clinic, 2016), which is theoretically grounded in the Competency Benchmarks (Fouad et al., 2009). Future research might benefit from use of such a measure; however, the TCE is in line with measures typically used in research looking at trainee competencies (i.e., locally adapted, theory-derived measures; Callahan & Watkins, 2018; Grus et al., 2016). Ideally, TCE ratings would also have been available for externs, who made up the majority of the sample, and baseline competency ratings would have been obtained to control for individual differences in competency and better assess growth.

Implications

Research

This study's findings demonstrate that supervision competence is a measurable construct, which research should continue to explore as both a predictor and an outcome. This study also demonstrates the utility of the PTESC in measuring supervision competence from the trainee perspective. Future research may enhance this instrument in several ways. First, after additional data collection, EFA could be used to reduce the number of items, reduce time burden on the rater, and enhance the PTESC's utility. Second, administering the measure to multiple raters (e.g., peer supervisors, self-assessment) would allow for further assessment of the measure's reliability and validity. Other new measures of supervision competence (e.g., the SE-SC

[Gonsalvez et al., 2017], GSAT [Hamilton et al., 2022]) could also be administered to assess convergent validity.

To expand our understanding of the influence of supervision competence on trainee development, future studies should use validated measures of trainee competence to supplement self-reporting. Trainee competencies might also be measured longitudinally over the course of the supervisory relationship. This study suggests that supervision competence may have varying degrees of influence on foundational versus functional competencies, which warrants further assessment. Such studies would also benefit from analyzing the influence of supervisor demographics. Finally, future research warrants creative and intentional study design to assess the impact of supervision competence on patient outcomes. Researchers may look to the examples of Bambling et al. (2006), Callahan et al. (2009), and Reese et al. (2009) for methodological inspiration in this area.

Theory

This study lends empirical support to the APA's (2014) seven domains of competency-based supervision. However, the Diversity domain stands out as being unique in nature and the most weakly related to supervision competence—although not for lack of importance (Hutman & Ellis, 2020; Inman & Ladany, 2014). Diversity competence, within supervision or otherwise, may be conceptually different from other competence domains in its greater importance placed on values and attitudes (e.g., valuing diversity, demonstrating cultural humility) than specific knowledge or skills. The multicultural orientation framework (Watkins et al., 2019) is a attitudes-additive perspective meant to complement knowledge- and skills-focused multicultural competence frameworks. It is worth considering how to better integrate a multicultural orientation framework into competency-based supervision and measures of supervision

competence, particularly as APA work groups consider updates to the PWCs and how to infuse diversity, equity, and inclusion principles throughout these expected competency areas.

Practice

The greatest implications for this study may be in the realm of practice. These findings support use of the PTESC, which can serve multiple functions. First, having trainees complete the PTESC in training settings provides necessary feedback to supervisors to inform and improve their implementation of competency-based supervision: it provides easy-to-understand scores across the seven domains with specific items that may elucidate areas of strength and weakness. Furthermore, the specificity and comprehensiveness of the PTESC may prove useful in facilitating dialogue between trainee and supervisor about areas for growth, which can be difficult feedback for trainees to provide. Creating a "feedback culture" in training programs has the potential to increase "uptake" of feedback and positively influence the supervisory relationship (Dudek et al., 2016). Of course, impression management and the inherent power dynamics of supervision may still prevent trainees from full transparency in such discussions; therefore, anonymous use of the PTESC for program evaluation (as is done at UCM) is another important practice. Anonymous feedback can be submitted and then aggregated to (a) provide supervisors with collective feedback that is less identifiable, and (b) examine trends in supervision competence and areas for improvement across a training program. Lastly, because supervision is itself one of the nine PWCs, the PTESC may prove useful in training of future supervisors: it could be administered by a supervisor observing the supervisor-in-training's work, or the supervisee of the supervisor-in-training, to provide feedback on their progress in attaining supervision competence.

Closing Thoughts

In closing, we must be able to assess supervision competence both for the sake of adapting to a competency-based training culture and to improve the quality of supervision. This study demonstrates that competent supervision enhances trainees' professional competencies, readying them to enter careers in health service psychology. Although more work is needed to understand how supervision competence influences patient outcomes, it seems likely that by improving trainees' competencies and the quality of clinical care they provide, patients should also benefit (Callahan & Watkins, 2018). Given supervision's duty to protect the public and its foundational role in training of health service psychologists, creating a culture of accountability for high quality, competent, and supportive supervision should be an area of utmost importance to the field.

References

American Psychological Association. (2002). *Book 1: Guidelines and principles for accreditation of programs in professional psychology*. Office of Program Consultation and Accreditation.

American Psychological Association. (2014). *Guidelines for clinical supervision in health service psychology*. https://www.apa.org/about/policy/guidelines-supervision.pdf

American Psychological Association. (2015). Standards of accreditation for health service psychology and accreditation operating procedures. https://www.apa.org/ed/accreditation/about/policies/standards-of-accreditation.pdf

American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct.* https://www.apa.org/ethics/code/

Association of State and Provincial Psychology Boards. (2019, August). *Supervision guidelines for education and training leading to licensure as a general applied provider.* https://cdn.ymaws.com/www.asppb.net/resource/resmgr/guidelines/supervision_guidelines_for_g.pdf

Bambling, M. (2014). Creating positive outcomes in clinical supervision. In C.E. Watkins, Jr. & D.L. Milne (Eds.), *The Wiley International Handbook of Clinical Supervision* (445-457). John Wiley & Sons, Ltd.

Bambling, M., King, R., Raue, P., Schweitzer, R., & Lambert, W. (2006). Clinical supervision: Its influence on client-rated working alliance and client symptom reduction in the brief treatment of major depression. *Psychotherapy Research, 16*(03), 317-331. https://doi.org/10.1080/10503300500268524

Bautista-Biddle, M. M., Pereira, L. M., & Williams, S. N. (2021). The fallacy of "good training experiences": The need to protect psychology trainees from harassment and the imperative of multiculturally competent supervision. *Training and Education in Professional Psychology, 15*(4), 323-330. http://dx.doi.org/10.1037/tep0000353

Bell, V., & O'Driscoll, C. (2018). The network structure of paranoia in the general population. *Social Psychiatry and Psychiatric Epidemiology*, *53*(7), 737-744. https://doi.org/10.1007/s00127-018-1487-0

Bernard, J. M., & Goodyear, R. K. (2014). *Fundamentals of Clinical Supervision* (5th ed.). Pearson.

Bogo, M., Regehr, C., Hughes, J., Power, R., & Globerman, J. (2002). Evaluating a measure of student field performance in direct service: Testing reliability and validity of explicit criteria. *Journal of Social Work Education, 38*(3), 385-401. https://doi.org/10.1080/10437797.2002.10779106

Bordin, E. S. (1983). A working alliance based model of supervision. *The Counseling Psychologist, 11*(1), 35-42.

Bradley, W. J., & Becker, K. D. (2021). Clinical supervision of mental health services: A systematic review of supervision characteristics and practices associated with formative and restorative outcomes. *The Clinical Supervisor*, *40*(1), 88-111. https://doi.org/10.1080/07325223.2021.1904312

Bradshaw, T., Butterworth, A., & Mairs, H. (2007). Does structured clinical supervision during psychosocial intervention education enhance outcome for mental health nurses and the service users they work with? *Journal of Psychiatric and Mental Health Nursing, 14*(1), 4-12. https://doi.org/10.1111/j.1365-2850.2007.01021.x

Callahan, J. L., Almstrom, C. M., Swift, J. K., Borja, S. E., & Heath, C. J. (2009). Exploring the contribution of supervisors to intervention outcomes. *Training and Education in Professional Psychology, 3*(2), 72. https://doi.org/10.1037/a0014294

Callahan, J. L., & Watkins Jr, C. E. (2018). The science of training III: Supervision, competency, and internship training. *Training and Education in Professional Psychology*, *12*(4), 245-261. https://doi.org/10.1037/tep0000208

Christensen, A. P., Gross, G. M., Golino, H. F., Silvia, P. J., & Kwapil, T. R. (2019). Exploratory graph analysis of the multidimensional schizotypy scale. *Schizophrenia Research*, *206*, 43-51. https://doi.org/10.1016/j.schres.2018.12.018

Christensen, A. P., Kennett, Y. N., Aste, T., Silvia, P. J., & Kwapil, T.R. (2018). Network structure of the Wisconsin Schizotypy Scales-Short Forms: Examining psychometric network filtering approaches. *Behavior Research Methods*, *50*(6), 2531-2550. https://doi.org/10.3758/s13428-018-1032-9

Constantine, M. G., & Sue, D. W. (2007). Perceptions of racial microaggressions among black supervisees in cross-racial dyads. *Journal of Counseling Psychology*, *54*(2), 142-153. https://doi.org/10.1037/0022-0167.54.2.142

Crawford, A. V., Green, S. B., Levy, R., Lo, W. J., Scott, L., Svetina, D., & Thompson, M. S. (2010). Evaluation of parallel analysis methods for determining the number of factors. *Educational and Psychological Measurement*, *70*(6), 885-901. https://doi.org/10.1177%2F0013164410379332

Crook Lyon, R., Heppler, A., Leavitt, L., & Fisher, L. (2008). Supervisory training experiences and overall supervisory development in predoctoral interns. *The Clinical Supervisor, 27*(2), 268-284. https://doi.org/10.1080/07325220802490877

Crook-Lyon, R. E., Presnell, J., Silva, L., Suyama, M., & Stickney, J. (2011). Emergent Supervisors: Comparing Counseling Center and Non-Counseling-Center Interns' Supervisory Training Experiences. *Journal of College Counseling*, *14*(1), 34-49. https://doi.org/10.1002/j.2161-1882.2011.tb00062.x

Davidson, R., & MacKinnon, J. G. (2001). Bootstrap tests: How many bootstraps?. *Econometric Reviews*, *19*(1), 55-68.

Dudek, N. L., Dojeiji, S., Day, K., & Varpio, L. (2016). Feedback to supervisors: is anonymity really so important?. *Academic Medicine*, *91*(9), 1305-1312. https://doi.org/10.1097/ACM.0000000000001170

Dunn, E. C., Richmond, T. K., Milliren, C. E., & Subramanian, S. V. (2015). Using cross classified multilevel models to disentangle school and neighborhood effects: An example focusing on smoking behaviors among adolescents in the United States. *Health Place, 31*, 224-232. https://doi.org/10.1016%2Fj.healthplace.2014.12.001

Ellis, M. V. (2010). Bridging the science and practice of clinical supervision: Some discoveries, some misconceptions. *The Clinical Supervisor*, *29*(1), 95-116. https://doi.org/10.1080/07325221003741910

Ellis, M. V. (2017). Narratives of harmful clinical supervision. *The Clinical Supervisor*, *36*(1), 20-87. https://doi.org/10.1080/07325223.2017.1297752

Ellis, M. V., D'Iuso, N., & Ladany, N. (2008). State of the art in the assessment, measurement, and evaluation of clinical supervision. In A.K. Hess, K.D. Hess, & T.H. Hess (Eds.), *Psychotherapy supervision: Theory, research, and practice* (pp. 473-499). John Wiley & Sons, Inc.

Ellis, M. V., & Ladany, N. (1997). Inferences concerning supervisees and clients in clinical supervision: An integrative review. In C. E. Watkins, Jr. (Ed.), *Handbook of psychotherapy supervision* (pp. 447–507). John Wiley & Sons Inc.

Ellis, M. V., Ladany, N., Krengel, M., & Schult, D. (1996). Clinical supervision research from 1981 to 1993: A methodological critique. *Journal of Counseling Psychology, 43*(1), 35. https://doi.org/10.1037/0022-0167.43.1.35

Ellis, M. V., Siembor, M. J., Swords, B. A., Morere, L., & Blanco, S. (2008, June). Prevalence and characteristics of harmful and inadequate clinical supervision. Paper presented at the 4th International Interdisciplinary Conference on Clinical Supervision, Buffalo, NY.

Epskamp, S., Cramer, A.O.J., Waldorp, L.J., Schmittmann, V.D., & Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *Journal of Statistical Software, 48*(4), 1-18. http://www.jstatsoft.org/v48/i04/.

Falender, C. A., Cornish, J. A. E., Goodyear, R., Hatcher, R., Kaslow, N. J., Leventhal, G., Shafranske, E., Sigmon, S. T., Stoltenberg, C., & Grus, C. (2004). Defining competencies in psychology supervision: A consensus statement. *Journal of Clinical Psychology, 60*(7), 771-785. https://doi.org/10.1002/jclp.20013

Falender, C., Grus, C., McCutcheon, S., Goodyear, R. K., Ellis, M., Doll, B., Miville, M., Rey-Casserly, C., & Kaslow, N. (2016). Guidelines for clinical supervision in health service psychology: Evidence and implementation strategies. *Psychotherapy Bulletin, 51*.

Falender, C. A., & Shafranske, E. P. (2004). *Clinical Supervision: A Competency-Based Approach*. American Psychological Association.

Falender, C. A., & Shafranske, E. P. (2017). Competency-based clinical supervision: Status, opportunities, tensions, and the future. *Australian Psychologist, 52*(2), 86-93. https://doi.org/10.1111/ap.12265

Falender, C. A., & Shafranske, E. P. (2021). *Clinical supervision: A Competency-based approach, second edition*. American Psychological Association.

Falender, C. A., Shafranske, E. P., & Ofek, A. (2014). Competent clinical supervision: Emerging effective practices. *Counselling Psychology Quarterly, 27*(4), 393-408. https://doi.org/10.1080/09515070.2014.934785

Finney, S. J., & DiStefano, C. (2006). Non-normal and categorical data in structural equation modeling. In G.R. Hancock & R.O. Mueller (Eds.), *Structural equation modeling: A second course*, 269-314. Information Age Publishing.

Fouad, N. A., Grus, C. L., Hatcher, R. L., Kaslow, N. J., Hutchings, P. S., Madson, M. B., Collins Jr, F. L., & Crossman, R. E. (2009). Competency benchmarks: A model for understanding and measuring competence in professional psychology across training levels. *Training and Education in Professional Psychology, 3*(4S), S5. https://doi.org/10.1037/a0015832

Freitas, G. J. (2002). The impact of psychotherapy supervision on client outcome: A critical examination of 2 decades of research. *Psychotherapy: Theory, Research, Practice, Training, 39*(4), 354. https://doi.org/10.1037/0033-3204.39.4.354

Gatmon, D., Jackson, D., Koshkarian, L., Martos-Perry, N., Molina, A., Patel, N., & Rodolfa, E. (2001). Exploring ethnic, gender, and sexual orientation variables in supervision: Do they really matter?. *Journal of Multicultural Counseling and Development*, *29*(2), 102-113. https://doi.org/10.1002/j.2161-1912.2001.tb00508.x

Gibson, A. S., Ellis, M. V., & Friedlander, M. L. (2019). Toward a nuanced understanding of nondisclosure in psychotherapy supervision. *Journal of Counseling Psychology, 66*(1), 114-121. http://dx.doi.org/10.1037/cou0000295

Golino, H., & Christensen, A. P. (2021). EGAnet: Exploratory Graph Analysis – A framework for estimating the number of dimensions in multivariate data using network psychometrics. R package version 0.9.8.

Golino, H. F., & Demetriou, A. (2017). Estimating the dimensionality of intelligence like data using Exploratory Graph Analysis. *Intelligence*, *62*, 54-70. http://dx.doi.org/10.1016/j.intell.2017.02.007

Golino, H. F., & Epskamp, S. (2017). Exploratory graph analysis: A new approach for estimating the number of dimensions in psychological research. *PloS one, 12*(6), e0174035. https://doi.org/10.1371/journal.pone.0174035

Golino, H., Shi, D., Christensen, A. P., Garrido, L. E., Nieto, M. D., Sadana, R., Thiyagarajan, J. A., & Martinez-Molina, A. (2020). Investigating the performance of exploratory graph analysis and traditional techniques to identify the number of latent factors: A simulation and tutorial. *Psychological Methods*, *25*(3), 292. https://dx.doi.org/10.1037/met0000255

Gonsalvez, C. J., & Crowe, T. P. (2014). Evaluation of psychology practitioner competence in clinical supervision. *American Journal of Psychotherapy*, *68*(2), 177-193.

Gonsalvez, C. J., & Freestone, J. (2007). Field supervisors' assessments of trainee performance: Are they reliable and valid?. *Australian Psychologist*, *42*(1), 23-32. https://doi.org/10.1080/00050060600827615

Gonsalvez, C. J., Hamid, G., Savage, N. M., & Livni, D. (2017). The supervision evaluation and supervisory competence scale: Psychometric validation. *Australian Psychologist, 52*(2), 94-103. https://doi.org/10.1111/ap.12269

Gonsalvez, C. J., Oades, L. G., & Freestone, J. (2002). The objectives approach to clinical supervision: Towards integration and empirical evaluation. *Australian Psychologist*, *37*(1), 68-77. https://doi.org/10.1080/00050060210001706706

Gonsalvez, C. J., Terry, J., Deane, F. P., Nasstasia, Y., Knight, R., & Gooi, C. H. (2021). End-of-placement failure rates among clinical psychology trainees: exceptional training and outstanding trainees or poor gate-keeping?. *Clinical Psychologist*, *25*(3), 294-305. https://doi.org/10.1080/13284207.2021.1927692

Goodyear, R. K., Murdock, N., Lichtenberg, J. W., McPherson, R., Koetting, K., & Petren, S. (2008). Stability and change in counseling psychologists' identities, roles, functions, and career satisfaction across 15 years. *The Counseling Psychologist*, *36*(2), 220-249. https://doi.org/10.1177%2F0011000007309481

Green, S.B., Redell, N., Thompson, M.S., & Levy, R. (2016). Accuracy of revised and traditional parallel analyses for assessing dimensionality with binary data. *Educational and Psychological Measurement, 76*(1), 5-21. https://doi.org/10.1177%2F0013164415581898

Gregus, S. J., Stevens, K. T., Seivert, N. P., Tucker, R. P., & Callahan, J. L. (2020). Student perceptions of multicultural training and program climate in clinical psychology doctoral programs. *Training and Education in Professional Psychology*, *14*(4), 293-307. https://doi.apa.org/doi/10.1037/tep0000289

Grus, C. L., Falender, C., Fouad, N. A., & Lavelle, A. K. (2016). A culture of competence: A survey of implementation of competency-based education and assessment. *Training and Education in Professional Psychology, 10*(4), 198-205. https://doi.org/10.1037/tep0000126

Hamilton, S. J., Briggs, L., Peterson, E. E., Slattery, M., & O'Donovan, A. (2022). Supporting conscious competency: Validation of the Generic Supervision Assessment Tool (GSAT). *Psychology and Psychotherapy: Theory, Research and Practice*, *95*(1), 113-136. https://doi.org/10.1111/papt.12369

Hayes, A. F. (2006). A primer on multilevel modeling. *Human Communication Research, 32*(4), 385-410. https://doi.org/10.1111/j.1468-2958.2006.00281.x

Hitzeman, C., Gonsalvez, C. J., Britt, E., & Moses, K. (2020). Clinical psychology trainees' self versus supervisor assessments of practitioner competencies. *Clinical Psychologist, 24*(1), 18-29. https://doi.org/10.1111/cp.12183

Hofmann, D.A. (1997). An overview of the logic and rationale of hierarchical linear models. *Journal of Management, 23*(6), 723-744. https://doi.org/10.1016/S0149-2063(97)90026-X

Holloway, E. L., & Neufeldt, S. A. (1995). Supervision: Its contributions to treatment efficacy. *Journal of Consulting and Clinical Psychology, 63*(2), 207-213. https://doi.org/10.1037/0022-006X.63.2.207

Hox, J.J. (2017). Cross-classified multilevel models. In J.J. Hox, M. Moerbeek, & R. Van de Schoot (Eds.), *Multilevel analysis: Techniques and applications* (pp. 161-172). Routledge. https://doi.org/10.4324/9781315650982

Hutman, H., & Ellis, M. V. (2020). Supervisee nondisclosure in clinical supervision: Cultural and relational considerations. *Training and Education in Professional Psychology, 14*(4), 308-315. http://dx.doi.org/10.1037/tep0000290

Inman, A. G., Hutman, H., Pendse, A., Devdas, L., Luu, L., & Ellis, M. V. (2014). Current trends concerning supervisors, supervisees, and clients in clinical supervision. In C. E. Watkins, Jr. & D.L. Milne (Eds.), *The Wiley international handbook of clinical supervision* (pp. 61-102). Wiley-Blackwell. https://doi.org/10.1002/9781118846360.ch4

Inman, A. G., & Ladany, N. (2014). Multicultural competencies in psychotherapy supervision. In F.T.L. Leong, L. Comas-Díaz, G. C. Nagayama Hall, V. C. McLoyd, & J. E. Trimble (Eds.), *APA Handbook of Multicultural Psychology, Vol. 2. Applications and Training* (pp. 643–658). American Psychological Association. https://doi.org/10.1037/14187-036

Jernigan, M. M., Green, C. E., Helms, J. E., Perez-Gualdron, L., & Henze, K. (2010). An examination of people of color supervision dyads: Racial identity matters as much as race. *Training and Education in Professional Psychology, 4*(1), 62-73. https://doi.org/10.1037/a0018110

Johnson, E. A., & Stewart, D. W. (2000). Clinical supervision in Canadian academic and service settings: The importance of education, training, and workplace support for supervisor development. *Canadian Psychology/Psychologie Canadienne, 41*(2), 124. https://doi.org/10.1037/h0086862

Kamen, C., Veilleux, J. C., Bangen, K. J., VanderVeen, J. W., & Klonoff, E. A. (2010). Climbing the stairway to competency: Trainee perspectives on competency development. *Training and Education in Professional Psychology, 4*(4), 227-234. https://doi.org/10.1037/a0021092

Kaslow, N. J., Borden, K. A., Collins Jr, F. L., Forrest, L., Illfelder-Kaye, J., Nelson, P. D., Rallo, J. S., Vasquez, M. J., & Willmuth, M. E. (2004). Competencies conference: Future directions in education and credentialing in professional psychology. *Journal of Clinical Psychology, 60*(7), 699-712. https://doi.org/10.1002/jclp.20016

Kavanagh, D. J., Spence, S. H., Strong, J., Wilson, J., Sturk, H., & Crow, N. (2003). Supervision practices in allied mental health: Relationships of supervision characteristics to perceived impact and job satisfaction. *Mental Health Services Research, 5*(4), 187–195. https://doi.org/1522-3434/03/1200-0187/0

Keith, T. Z., Caemmerer, J. M., & Reynolds, M. R. (2016). Comparison of methods for factor extraction for cognitive test-like data: Which overfactor, which underfactor?. *Intelligence, 54*, 37-54. https://doi.org/10.1016/j.intell.2015.11.003

Kim, E. S., Dedrick, R. F., Cao, C., & Ferron, J. M. (2016). Multilevel factor analysis: Reporting guidelines and a review of reporting practices. *Multivariate Behavioral Research, 51*(6), 881-898. https://doi.org/10.1080/00273171.2016.1228042

Ladany, N., Mori, Y., & Mehr, K. E. (2013). Effective and ineffective supervision. *The Counseling Psychologist, 41*(1), 28-47. https://doi.org/10.1177/0011000012442648

Larkin, K. T., & Morris, T. L. (2015). The process of competency acquisition during doctoral training. *Training and Education in Professional Psychology, 9*(4), 300-308. https://doi.org/10.1037/tep0000091

Leckie, G. (2013). Cross-classified multilevel models – Concepts. LEMMA VLE Module 12, 1-60. http://www.bristol.ac.uk/cmm/learning/course.html

Lehrman-Waterman, D., & Ladany, N. (2001). Development and validation of the Evaluation Process Within Supervision Inventory. *Journal of Counseling Psychology, 48*(2), 168. https://doi.org/10.1037/0022-0167.48.2.168

Little, T. D., Rhemtulla, M., Gibson, K., & Schoemann, A. M. (2013). Why the items versus parcels controversy needn't be one. *Psychological Methods*, *18*(3), 285-300. https://doi.org/10.1037/a0033266

Lucock, M. P., Hall, P., & Noble, R. (2006). A survey of influences on the practice of psychotherapists and clinical psychologists in training in the UK. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice, 13*(2), 123-130. https://doi.org/10.1002/cpp.483

Magnuson, S., Wilcoxon, S. A., & Norem, K. (2000). A profile of lousy supervision: Experienced counselors' perspectives. *Counselor Education and Supervision*, *39*(3), 189-202. https://doi.org/10.1002/j.1556-6978.2000.tb01231.x

McNamara, M. L., Kangos, K. A., Corp, D. A., Ellis, M. V., & Taylor, E. J. (2017). Narratives of harmful clinical supervision: Synthesis and recommendations. *The Clinical Supervisor*, *36*(1), 124-144. https://doi.org/10.1080/07325223.2017.1298488

Mehr, K. E., & Daltry, R. M. (2022). Supervisor self-disclosure, the supervisory alliance, and trainee willingness to disclose. *Professional Psychology: Research and Practice*, *53*(3), 313–317. https://doi.org/10.1037/pro0000424

Mehr, K. E., Ladany, N., & Caskie, G. I. (2015). Factors influencing trainee willingness to disclose in supervision. *Training and Education in Professional Psychology, 9*(1), 44-51. https://doi.org/10.1037/tep0000028

Milne, D. (2009). *Evidence-based clinical supervision: Principles and Practice.* BPS Blackwell.

Milne, D., Aylott, H., Fitzpatrick, H., & Ellis, M. V. (2008). How does clinical supervision work? Using a "best evidence synthesis" approach to construct a basic model of supervision. *The Clinical Supervisor, 27*(2), 170-190. https://doi.org/10.1080/07325220802487915

Milne, D. L., & James, I. A. (2002). The observed impact of training on competence in clinical supervision. *British Journal of Clinical Psychology, 41*(1), 55-72. https://doi.org/10.1348/014466502163796

Milne, D. L., Reiser, R. P., Cliffe, T., & Raine, R. (2011). SAGE: preliminary evaluation of an instrument for observing competence in CBT supervision. *The Cognitive Behaviour Therapist, 4*(4), 123-138. https://doi.org/10.1017/S1754470X11000079

Milne, D.L., Sheikh, A.I., Pattison, S., & Wilkinson, A. (2011). Evidence-based training for clinical supervisiors: A systematic review of 11 controlled studies. *The Clinical Supervisor, 30*(1), 53-71, https://doi.org/10.1080/07325223.2011.564955

Murphy-Shigematsu, S. (2010). Microaggressions by supervisors of color. *Training and Education in Professional Psychology, 4*(1), 16-18. https://doi.org/10.1037/a0017472

Newman, D. S., Hazel, C. E., Faler, A., & Pomerantz, L. (2021). Supervision training in APA-accredited school psychology doctoral programs: An analysis of syllabi. *Training and Education in Professional Psychology, 15*(3), 176-188. https://doi.org/10.1037/tep0000310

Norcross, J. C., & Rogan, J. D. (2013). Psychologists conducting psychotherapy in 2012: Current practices and historical trends among Division 29 members. *Psychotherapy, 50*(4), 490-495. https://doi.org/10.1037/a0033512

Orlinsky, D. E., Ronnestad, M. H., & Collaborative Research Network of the Society for Psychotherapy Research. (2005). *How Psychotherapists Develop: A Study of Therapeutic Work and Professional Growth*. American Psychological Association.

Park, E. H., Ha, G., Lee, S., Lee, Y. Y., & Lee, S. M. (2019). Relationship between the supervisory working alliance and outcomes: A meta-analysis. *Journal of Counseling & Development, 97*(4), 437-446. https://doi.org/10.1002/jcad.12292

Pattengale, N. D., Alipour, M., Bininda-Emonds, O. R., Moret, B. M., & Stamatakis, A. (2010). How many bootstrap replicates are necessary?. *Journal of Computational Biology, 17*(3), 337-354. https://doi.org/10.1089/cmb.2009.0179

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). *Making sense of factor analysis: The use of factor analysis for instrument development in health care research.* Sage.

Phillips, J. C., Parent, M. C., Dozier, V. C., & Jackson, P. L. (2017). Depth of discussion of multicultural identities in supervision and supervisory outcomes. *Counselling Psychology Quarterly*, *30*(2), 188-210. https://doi.org/10.1080/09515070.2016.1169995

Price, S. D., Callahan, J. L., & Cox, R. J. (2017). Psychometric investigation of competency benchmarks. *Training and Education in Professional Psychology, 11*(3), 128-139. https://doi.org/10.1037/tep0000133

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Reese, R. J., Usher, E. L., Bowman, D. C., Norsworthy, L. A., Halstead, J. L., Rowlands, S. R., & Chisholm, R. R. (2009). Using client feedback in psychotherapy training: An analysis of its influence on supervision and counselor self-efficacy. *Training and Education in Professional Psychology, 3*(3), 157. https://doi.org/10.1037/a0015673

Reiser, R. P., Cliffe, T., & Milne, D. L. (2018). An improved competence rating scale for CBT Supervision: Short-SAGE. *The Cognitive Behaviour Therapist*, *11*(e7), 1-16. https://doi.org/10.1017/S1754470X18000065

Reiser, R. P., & Milne, D. L. (2014). A systematic review and reformulation of outcome evaluation in clinical supervision: Applying the fidelity framework. *Training and Education in Professional Psychology, 8*(3), 149. https://doi.org/10.1037/tep0000031

Roberts, M. C., Borden, K. A., Christiansen, M. D., & Lopez, S. J. (2005). Fostering a Culture Shift: Assessment of Competence in the Education and Careers of Professional Psychologists. *Professional Psychology: Research and Practice*, *36*(4), 355-361. https://doi.org/10.1037/0735-7028.36.4.355

Rosseel, Y. (2012). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1-36. https://www.jstatsoft.org/v48/i02/

Roth, A. D., Pilling, S., & Turner, J. (2010). Therapist training and supervision in clinical trials: Implications for clinical practice. *Behavioural and Cognitive Psychotherapy, 38*(3), 291-302. https://doi.org/10.1017/S1352465810000068

Rousmaniere, T. G., Swift, J. K., Babins-Wagner, R., Whipple, J. L., & Berzins, S. (2016). Supervisor variance in psychotherapy outcome in routine practice. *Psychotherapy Research*, *26*(2), 196-205. https://doi.org/10.1080/10503307.2014.963730

Ruscio, J., & Roche, B. (2012). Determining the number of factors to retain in an exploratory factor analysis using comparison data of known factorial structure. *Psychological Assessment, 24*(2), 282-292. https://doi.org/10.1027/1614-2241/a000034

Schoenwald, S. K., Sheidow, A. J., & Chapman, J. E. (2009). Clinical supervision in treatment transport: Effects on adherence and outcomes. *Journal of Consulting and Clinical Psychology, 77*(3), 410–421. https://doi.org/10.1037/a0013788

Scott, K. J., Ingram, K. M., Vitanza, S. A., & Smith, N. G. (2000). Training in supervision: A survey of current practices. *The Counseling Psychologist, 28*(3), 403-422. https://doi.org/10.1177/0011000000283007

Scott, T. L., Pachana, N. A., & Sofronoff, K. (2011). Survey of current curriculum practices within Australian postgraduate clinical training programmes: Students' and programme directors' perspectives. *Australian Psychologist*, *46*(2), 77-89. https://doi.org.10.1111/j.1742-9544.2011.00030.x

Soheilian, S. S., Inman, A. G., Klinger, R. S., Isenberg, D. S., & Kulp, L. E. (2014). Multicultural supervision: Supervisees' reflections on culturally competent supervision. *Counselling Psychology Quarterly*, *27*(4), 379-392. https://doi.org/10.1080/09515070.2014.961408

Stedman, J. M., Schoenfeld, L. S., & O'Donnell, L. (2013). An investigation of internship directors' perspectives on the learning objectives required by the commission on accreditation. *Training and Education in Professional Psychology, 7*(2), 134-138. https://doi.org/10.1037/a0031660

Strom, T. Q., Arbisi, P., Thuras, P. D., Leskela, J., Finn, J., Anderson, C., & Siegel, W. (2016). Summative intern performance evaluations: A wealth of information or a gathering of "two cents"? *Training and Education in Professional Psychology, 10*(2), 78-83. https://doi.org/10.1037/tep0000114

Sue, D. W., Capodilupo, C. M., Torino, G. C., Bucceri, J. M., Holder, A., Nadal, K. L., & Esquilin, M. (2007). Racial microaggressions in everyday life: implications for clinical practice. *American Psychologist, 62*(4), 271-286. https://doi.org/10.1037/0003-066X.62.4.271

Tabachnick, B.G. & Fidell, L.S. (2007). Principal components and factor analysis. In *Using Multivariate Statistics* (pp. 607-675). Pearson.

Tracey, T. J. G., Wampold, B. E., Lichtenberg, J. W., & Goodyear, R. K. (2014). Expertise in psychotherapy: An elusive goal? *American Psychologist, 69*(3), 218-229. https://doi.org/10.1037/a0035099

University of North Texas Psychology Clinic Executive Committee. (2016). *Practicum Evaluation Form*. Denton, TX: Author.

Vas, S.N., Dave, P., & Kass, A. (2021). *The University of Chicago Medicine Department of Psychiatry and Behavioral Neuroscience Psychology Trainee Evaluation of Supervision Competencies*. In C.A. Falender & E.P. Shafranske (Eds.), *Clinical supervision: A Competency-based approach, second edition.* American Psychological Association.

Watkins Jr, C. E. (2011). Does psychotherapy supervision contribute to patient outcomes? Considering thirty years of research. *The Clinical Supervisor, 30*(2), 235-256. https://doi.org/10.1080/07325223.2011.619417

Watkins Jr, C. E. (2014). The supervisory alliance: A half century of theory, practice, and research in critical perspective. *American Journal of Psychotherapy, 68*(1), 19-55.

Watkins Jr, C. E. (2017). How does psychotherapy supervision work? Contributions of connection, conception, allegiance, alignment, and action. *Journal of Psychotherapy Integration, 27*(2), 201. https://doi.org/10.1037/int0000058

Watkins Jr, C. E., Hook, J. N., Owen, J., DeBlaere, C., Davis, D. E., & Van Tongeren, D. R. (2019). Multicultural orientation in psychotherapy supervision: Cultural humility, cultural comfort, and cultural opportunities. *American Journal of Psychotherapy, 72*(2), 38-46. https://doi.org/10.1176/appi.psychotherapy.20180040

Wheeler, S., & Barkham, M. (2014). A core evaluation battery for supervision. In In C.E. Watkins, Jr. & D.L. Milne (Eds.), *The Wiley International Handbook of Clinical Supervision* (pp. 365-385). John Wiley & Sons, Ltd.

Wheeler, S., & Richards, K. (2007). The impact of clinical supervision on counsellors and therapists, their practice and their clients. A systematic review of the literature. *Counselling and Psychotherapy Research, 7*(1), 54-65. https://doi.org/10.1080/14733140601185274

White, E., & Winstanley, J. (2010). A randomised controlled trial of clinical supervision: Selected findings from a novel Australian attempt to establish the evidence base for causal relationships with quality of care and patient outcomes, as an informed contribution to mental health nursing practice development. *Journal of Research in Nursing, 15*(2), 151-167. https://doi.org/10.1177/1744987109357816

Wrape, E. R., Callahan, J. L., Ruggero, C. J., & Watkins Jr, C. E. (2015). An exploration of faculty supervisor variables and their impact on client outcomes. *Training and Education in Professional Psychology, 9*(1), 35. https://doi.org/10.1037/tep0000014

Appendix A

The Psychology Trainee Evaluation of Supervision Competencies (PTESC)

Confidential

UCM Supervisor Evaluation 2018-19

Please complete this evaluation of your supervisor.

Thank you!

What is today's date?	_____ (MM/DD/YEAR)
What is your name?	_____
What is your position?	◯ Adult/Health Intern ◯ Adult Neuropsychology Intern ◯ Child/Pediatric Intern ◯ Pediatric Neuropsychology Intern ◯ Psychology Extern: Adult Section ◯ Psychology Extern: Child Section ◯ Psychology Postdoctoral Fellow ◯ Social Work Trainee ◯ Other
Other position:	_____
What is your supervisor's name? (List your supervisor's first and last name without titles or degrees, e.g., Jane Doe)	_____
In what capacity are you evaluating this supervisor? Check all that apply.	☐ Internship Outpatient Psychotherapy Supervisor ☐ Internship Rotation Supervisor ☐ Externship Psychotherapy Supervisor ☐ Externship Assessment Supervisor ☐ Group Supervision ☐ Peer Supervision ☐ Mentor ☐ Training Director ☐ Fellowship Supervisor ☐ Social Work Supervisor ☐ Other
Other capacity:	_____
At what time point in the training program are you evaluating this supervisor?	◯ 3 months ◯ 6 months ◯ 9 months ◯ 12 months ◯ 18 months ◯ 24 months ◯ 36 months ◯ Other
Other time point:	_____

Confidential

DOMAIN A: SUPERVISOR COMPETENCE

Goal A1: Assurance of competence in the provision of services

Provides competent supervision of services to ensure welfare of patients.

- ◯ Behavior Never Displayed/Observed
- ◯ Behavior Rarely Displayed
- ◯ Behavior Frequently Displayed
- ◯ Behavior Typically Displayed
- ◯ Behavior (Almost) Always Displayed
- ◯ Not Applicable

Confidential

Please provide a rating for each specific objective listed below:						
	Behavior Never Displayed	Behavior Rarely Displayed	Behavior Frequently Displayed	Behavior Typically Displayed	Behavior (Almost) Always Displayed	Not Applicable
Ensures that patients receive competent services and protects others from harm.	○	○	○	○	○	○
Possesses and demonstrates up-to-date knowledge and skill about the areas being supervised.	○	○	○	○	○	○
Takes reasonable steps to ensure competence when less familiar with the areas being supervised.	○	○	○	○	○	○
Sets appropriate boundaries and seeks consultation when supervisory issues are outside domain of supervisory competence.	○	○	○	○	○	○
Refers trainees to other resources (e.g., consultation, research, etc.) when appropriate.	○	○	○	○	○	○
Demonstrates knowledge about the context of supervision (e.g., expectations of the system within which the trainee works, departmental/institutional policies, etc.).	○	○	○	○	○	○
Demostrates knowledge about relevant events that may impact patient care (e.g., billing and administrative procedures, etc.) in the organizational context.	○	○	○	○	○	○
Consistently enforces appropriate standards for billing procedures, documentation, and administrative protocols, and encourages trainees to become fluent in this domain.	○	○	○	○	○	○
Demonstrates flexibility in teaching modalities, case conceptualization, and treatment plan suggestions.	○	○	○	○	○	○

Confidential

Demonstrates scientific thinking and appropriate translation of scientific findings to practice.	○	○	○	○	○	○
Collaborates with all faculty/staff involved in the training process at the site.	○	○	○	○	○	○
Communicates with trainee's graduate program as necessary, especially when performance problems need to be addressed.	○	○	○	○	○	○
Strives for diversity competence across populations and settings.	○	○	○	○	○	○
Demonstrates knowledge about diversity issues that are specifically relevant to the setting and environment within which the trainee works.	○	○	○	○	○	○
Possesses relevant knowledge, skills, and values/attitudes to provide culturally sensitive care and supervision.	○	○	○	○	○	○
Provides evaluative feedback routinely to enhance development of trainee competence.	○	○	○	○	○	○
Demonstrates and models transparency in the process of communication and encourages similar behavior in supervisee	○	○	○	○	○	○
Strives to be competent in the use of technology in clinical care (including distance supervision).	○	○	○	○	○	○
Demonstrates awareness of the policies and procedures in place for ethical practice of telepsychology, social media, and electronic communication.	○	○	○	○	○	○
Possesses relevant knowledge about legal issues sepcific to technology, supervision, and practice.	○	○	○	○	○	○

Confidential

Models ethical practice, decision-making, and professionalism by facilitating thoughful discussion regarding relevant issues (e.g., social networking).	○	○	○	○	○	○

Goal A2: Competence in the provision of supervision

Demonstrates competence in the provision of supervision.

○ Behavior Never Displayed/Observed
○ Behavior Rarely Displayed
○ Behavior Frequently Displayed
○ Behavior Typically Displayed
○ Behavior (Almost) Always Displayed
○ Not Applicable

Please provide a rating for each specific objective listed below.

	Behavior Never Displayed	Behavior Rarely Displayed	Behavior Frequently Displayed	Behavior Typically Displayed	Behavior (Almost) Always Displayed	Not Applicable
Seeks to attain and maintain competence in the practice of supervision through consultation, education, and training.	○	○	○	○	○	○
Demonstrates requisite knowledge of models, theories, modalities, and research on clinical supervision and relevant skills.	○	○	○	○	○	○
Demonstrates committment to knowing and utilizing available psychological sciece related to supervision	○	○	○	○	○	○
Manages supervisory relationship while enhancing trainee's skills.	○	○	○	○	○	○
Demonstrates familiarity with and utilizes a developmental approach to supervision	○	○	○	○	○	○
Formally and/or informally assesses the learning needs and developmental level of the trainee on an ongoing basis	○	○	○	○	○	○
Continually adjusts teaching model to skill level in accordance with the developmental model of supervision	○	○	○	○	○	○

Confidential

Provides input consistent with developmental needs of trainee (e.g., less specific feedback over time, increased encouragement of higher level case conceptualization, promotion of autonomous thinking appropriate to level of trainee, etc.).	○	○	○	○	○	○

DOMAIN B: DIVERSITY

Goal: Diversity Competence

Strives to develop diversity competence as an element of supervision competence, and infuses diversity into all aspects of clinical practice and research.

- ○ Behavior Never Displayed/Observed
- ○ Behavior Rarely Displayed
- ○ Behavior Frequently Displayed
- ○ Behavior Typically Displayed
- ○ Behavior (Almost) Always Displayed
- ○ Not Applicable

Confidential

Please provide a rating for each specific objective listed below.						
	Behavior Never Displayed	Behavior Rarely Displayed	Behavior Frequently Displayed	Behavior Typically Displayed	Behavior (Almost) Always Displayed	Not Applicable
Demonstrates awareness and knowledge of diversity in all of its forms.	○	○	○	○	○	○
Develops and maintains self-awareness regarding own diversity competence, which incudes attitudes, knowledge, and skills.	○	○	○	○	○	○
Demonstrates efforts to be introspective, revise and update knowledge, and advance diversity skills.	○	○	○	○	○	○
Models openness to self-exploration, understanding of one's own biases, and willingness to pursue education or consultation as necessary.	○	○	○	○	○	○
Recognizes the value of and pursues ongoing training in diversity competence as part of professional development and life-long learning	○	○	○	○	○	○
Attempts to be knowledgeable about the effects of bias and prejudice, and as necessary, models advocacy behaviors to promote change.	○	○	○	○	○	○
Serves as a role model regarding diversity, knowledge, skills, and attitudes.	○	○	○	○	○	○
Strives to be familiar with the literature concerning diversity competence in supervision.	○	○	○	○	○	○
Encourages sensitivity to diversity in all its forms.	○	○	○	○	○	○
Establishes a respectful supervisory relationship to facilitate diversity competence.	○	○	○	○	○	○

Confidential

Creates a safe environment within which to address diversity issues in clinical care, supervision, and organizational context.	○	○	○	○	○	○
Manages individual difference variables that may impact the supervisory relationship.	○	○	○	○	○	○
Assists with and encourages the development of cogent case formulation that includes individual difference variables.	○	○	○	○	○	○
Assists trainee in navigating tension between personal and professional values in providing competent patient care.	○	○	○	○	○	○

DOMAIN C: SUPERVISORY RELATIONSHIP

Goal: Development and Maintenance of a Positive Supervisory Relationship

Creates a supervisory relationship that facilitates effective clinical supervision.

- ○ Behavior Never Displayed/Observed
- ○ Behavior Rarely Displayed
- ○ Behavior Frequently Displayed
- ○ Behavior Typically Displayed
- ○ Behavior (Almost) Always Displayed
- ○ Not Applicable

Confidential

Please provide a rating for each specific objective listed below.

	Behavior Never Displayed	Behavior Rarely Displayed	Behavior Frequently Displayed	Behavior Typically Displayed	Behavior (Almost) Always Displayed	Not Applicable
Values, creates, and maintains a collaborative relationship that promotes the supervisee's competence.	○	○	○	○	○	○
Specifies responsibilities and expectations of both parties in the supervisory relationship.	○	○	○	○	○	○
Identifies expected program competencies and performance standards.	○	○	○	○	○	○
Collaboratively develops individualized goals for supervision in the form of a clearly specified supervisory contract.	○	○	○	○	○	○
Collaboratively assesses progress towards goals on an ongoing basis.	○	○	○	○	○	○
Regularly reviews progress of trainee and the effectiveness of the supervisory relationship and addresses relevant issues as necessary.	○	○	○	○	○	○
Demonstrates sensitivity to multiple potential roles with the trainee and exhibits ability to perform and balance multiple roles.	○	○	○	○	○	○
Promotes growth and self-assessment in the trainee.	○	○	○	○	○	○
Encourages and uses evaluative feedback from the trainee on an ongoing basis.	○	○	○	○	○	○
Demonstrates respect for trainees, patients, and colleagues.	○	○	○	○	○	○
Promotes autonomy appropriate to supervisee's level of training.	○	○	○	○	○	○

Confidential

DOMAIN D: PROFESSIONALISM

Goal: Comportment reflecting the fundamental values of professional psychology

Prioritizes needs and welfare of patients and trainees, and exhibits integrity, professional behavior, accountability, and concern for the value of others.

- ◯ Behavior Never Displayed/Observed
- ◯ Behavior Rarely Displayed
- ◯ Behavior Frequently Displayed
- ◯ Behavior Typically Displayed
- ◯ Behavior (Almost) Always Displayed
- ◯ Not Applicable

Please provide a rating for each specific objective listed below.

	Behavior Never Displayed	Behavior Rarely Displayed	Behavior Frequently Displayed	Behavior Typically Displayed	Behavior (Almost) Always Displayed	Not Applicable
Models professionalism through own behavior and interactions with others.	◯	◯	◯	◯	◯	◯
Teaches knowledge, skills, and attitudes associated with professionalism.	◯	◯	◯	◯	◯	◯
Provides ongoing feedback and evaluation of trainee progress towards meeting professional expectations appropriate for level of education and training.	◯	◯	◯	◯	◯	◯
Is available as needed for supervision/consultation.	◯	◯	◯	◯	◯	◯
Provides own work samples to illustrate specific issues.	◯	◯	◯	◯	◯	◯
Sets and keeps regularly scheduled meeting times.	◯	◯	◯	◯	◯	◯
Provides for a covering supervisor during absences.	◯	◯	◯	◯	◯	◯

DOMAIN E: ASSESSMENT, EVALUATION, & FEEDBACK

Goal: Provision of appropriate and timely evaluation and feedback

Provides appropriate and timely assessment, evaluation, and feedback.

- ◯ Behavior Never Displayed/Observed
- ◯ Behavior Rarely Displayed
- ◯ Behavior Frequently Displayed
- ◯ Behavior Typically Displayed
- ◯ Behavior (Almost) Always Displayed
- ◯ Not Applicable

Confidential

Please provide a rating for each specific objective listed below.						
	Behavior Never Displayed	Behavior Rarely Displayed	Behavior Frequently Displayed	Behavior Typically Displayed	Behavior (Almost) Always Displayed	Not Applicable
Promotes openness and transparency in assessment and feedback by relating this information to competency development.	○	○	○	○	○	○
Describes how supervision is to be conducted and follows model described.	○	○	○	○	○	○
Utilizes multiple methods of evaluation (e.g., live observation, chart review, audio/video review) to monitor performance.	○	○	○	○	○	○
Provides direct, clear, timely, and behaviorally anchored feedback.	○	○	○	○	○	○
Is mindful of the impact of feedback on the supervisory relationshp.	○	○	○	○	○	○
Incorporates trainee self-assessment into the evaluation process.	○	○	○	○	○	○
Highlights trainee strengths and impact on performance.	○	○	○	○	○	○
Seeks feedback from trainee about supervision and incorporates this information appropriately.	○	○	○	○	○	○
Provides effective formative and summative feedback.	○	○	○	○	○	○
Deonstrates knowledge of evaluation, process, and outcomes.	○	○	○	○	○	○
Observes both positive and negative trainee behaviors.	○	○	○	○	○	○
Balances between being supportive and challenging.	○	○	○	○	○	○

Confidential

Written material (e.g., notes, reports) is reviewed and returned with appropriate feedback in a timely manner.	○	○	○	○	○	○
Supervisor submits all materials in accordance with departmental deadlines and policies.	○	○	○	○	○	○

DOMAIN F: TRAINEE REMEDIATION & MANAGEMENT OF PROBLEMS OF PROFESSIONAL COMPETENCE

Goal: Management of problems of professional competence

Addresses problems with professional competence and provides support/remediation as necessary.

- ○ Behavior Never Displayed/Observed
- ○ Behavior Rarely Displayed
- ○ Behavior Frequently Displayed
- ○ Behavior Typically Displayed
- ○ Behavior (Almost) Always Displayed
- ○ Not Applicable

Please enter a rating for each specific objective listed below.

	Behavior Never Displayed	Behavior Rarely Displayed	Behavior Frequently Displayed	Behavior Typically Displayed	Behavior (Almost) Always Displayed	Not Applicable
Understands and adheres to the supervisory contract and procedures related to performance evaluations.	○	○	○	○	○	○
Identifies current or potential problems of professional competence promptly and directly communicates them to the supervisee.	○	○	○	○	○	○
Addresses problems in a timely manner to allow opportunities for change.	○	○	○	○	○	○
Develops and implements an appropriate support/remediation plan for problems of professional competence.	○	○	○	○	○	○
Communicates with supervisee's graduate program as necessary.	○	○	○	○	○	○
Takes ethically appropriate action in response to supervisee's performance problems.	○	○	○	○	○	○

Confidential

DOMAIN G: ETHICAL, LEGAL, & REGULATORY CONSIDERATIONS

Goal: Ethical and Appropriate Conduct

Values and models ethical behavior and adheres to relevant legal and regulatory parameters

- ◯ Behavior Never Displayed/Observed
- ◯ Behavior Rarely Displayed
- ◯ Behavior Frequently Displayed
- ◯ Behavior Typically Displayed
- ◯ Behavior (Almost) Always Displayed
- ◯ Not Applicable

Please provide a rating for each specific objective listed below.

	Behavior Never Displayed	Behavior Rarely Displayed	Behavior Frequently Displayed	Behavior Typically Displayed	Behavior (Almost) Always Displayed	Not Applicable
Demonstrates knowledge of ethics and legal issues specific to supervision.	◯	◯	◯	◯	◯	◯
Demonstrates knowledge of and upholds professional ethical standards, and encourages this practice among supervisees.	◯	◯	◯	◯	◯	◯
Models ethical practice and decision-making and conducts self in accordance with APA and other guidelines and laws/regulations.	◯	◯	◯	◯	◯	◯
Upholds primary ethical and legal obligation to protect the welfare of the patient	◯	◯	◯	◯	◯	◯
Provides information about expectations for and parameters of supervision in a clearly specified contractual agreement.	◯	◯	◯	◯	◯	◯
Maintains accurate and timely documentation of trainee performance related to expectations for comptency and professional development.	◯	◯	◯	◯	◯	◯
Manages responsibility as "gatekeeper" to the profession by assessing suitability to enter and remain in the field.	◯	◯	◯	◯	◯	◯

ACQUISITION OF PROFESSION WIDE COMPETENCIES

Confidential

Please indicate your agreement with each of the statements listed below.					
	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
My training experience with this supervisor facilitated the acquisition of competency in science, research, and evaluation.	○	○	○	○	○
My training experience with this supervisor facilitated the acquisition of competency in ethics and legal issues.	○	○	○	○	○
My training experience with this supervisor facilitated the acquisition of competency in individual and cultural diversity.	○	○	○	○	○
My training experience with this supervisor facilitated the acquisition of competency in professional values, attitudes, and behaviors.	○	○	○	○	○
My training experience with this supervisor facilitated the acquisition of competency in communication and interpersonal skills.	○	○	○	○	○
My training experience with this supervisor facilitated the acquisition of competency in psychological assessment and diagnosis.	○	○	○	○	○
My training experience with this supervisor facilitated the acquisition of competency in psychotherapeutic intervention.	○	○	○	○	○
My training experience with this supervisor facilitated the acquisition of competency in supervision, education, and training.	○	○	○	○	○
My training experience with this supervisor facilitated the acquisition of competency in consultation and interdisciplinary practice.	○	○	○	○	○

Appendix B

Psychology Trainee Competency Evaluation (TCE)

THE UNIVERSITY OF CHICAGO MEDICINE
AT THE FOREFRONT OF MEDICINE®

DEPARTMENT OF PSYCHIATRY & BEHAVIORAL NEUROSCIENCE
Psychology Trainee Competency Evaluation
(Vas, Dave, & Kass, 2015)

Trainee:
Supervisor:
Date:
Experience Being Evaluated:

This evaluation is based on the following sources of information: (check all that apply)
☐ Direct observation ☐Discussions in supervision ☐Audio/video tape review
☐ Feedback from others ☐Participation in meetings ☐ Review of clinical records
☐ Other (specify)_____
Insert at least one date that direct observation or video tape review occurred: 9/10/2019

Please use the following rating scale for evaluation for outpatient clinics and rotations. For each area of competency, general goals and several specific objectives are listed. **Please provide a rating for ALL items (i.e. Goals AND Objectives), unless the item is rotation-specific.** Space is provided at the end of the form for narrative description of the trainee's level of functioning. Please remember that all ratings should be made relative to the level of performance expected given the point of the year at which the evaluation is conducted.

COMPETENCY RATING DESCRIPTIONS		
NA/NO	**NA/NO**	**Not Applicable for this training experience OR Not Observed during this training experience.**
5	**A**	**Advanced/Skills comparable to autonomous practice at the licensure level.** Rating expected at completion of postdoctoral training. Competency attained at full psychology staff privilege level, however as an unlicensed trainee, supervision is required while in training status.
4	**HI**	**High Intermediate/Occasional supervision needed.** A frequent rating at completion of internship. Competency attained in all but non-routine cases; supervisor provides overall management of trainee's activities; depth of supervision varies as clinical needs warrant.
3	**I**	**Intermediate/Should remain a focus of supervision.** Common rating throughout internship and externship. Routine supervision of each activity.
2	**E**	**Entry level/Continued intensive supervision is needed.** Most common rating for externship. Routine, but intensive, supervision is needed.
1	**R**	**Needs remediation.** Requires remedial work if trainee is on internship or externship.

Competency A: SCIENCE, RESEARCH, AND EVALUATION

A1. Goal: Acquisition and Utilization of Current Scientific Knowledge

	Demonstrates necessary self-direction in gathering clinical and research information to practice independently and competently.

Specific Objectives:

	Demonstrates commitment to evidence-based practice that integrates the best available research with clinical expertise in the context of patient characteristics, culture, and preferences.
	Independently seeks out current scientific information to enhance clinical practice and other relevant areas by utilizing available databases, professional literature, seminars, training, and other resources.
	Critically evaluates health and behavior research relevant to populations to be served.
	Recognizes limits to competence and areas of expertise and takes steps to address these issues.
	Requests and utilizes supervisor's suggestions of additional information and resources.
	Demonstrates motivation to increase knowledge and expand range of professional skills through reading and supervision/consultation as necessary.

A2. Goal: Program/Outcome Evaluation

	Demonstrates appropriate knowledge and use of program/outcome evaluation.

Specific Objectives:

	Use of research skills for program development and evaluation as well as for quality improvement related to health care services.
	Uses appropriate measures to routinely evaluate outcomes as necessary.
	Uses evaluation data appropriately to guide further decisions and change process and/or outcomes.
	Provides evaluative feedback to patients, supervisors, and colleagues as necessary in order to improve process and/or outcomes.
	Seeks supervision/consultation as necessary to enhance competence in program/outcome evaluation.

A3. Goal: Conducting and Disseminating Scientific Research

	Demonstrates competence in the various elements of the process of conducting and disseminating scientific research.

Specific Objectives:

	Demonstrates familiarity with empiricism and health research methods.
	Conducts research that contributes to the scientific and professional knowledge base or evaluates the effectiveness of various professional activities in health care and health promotion.
	Collaborates with faculty supervisor/mentor regarding efforts to disseminate results (e.g., manuscript preparation) in an appropriate and timely manner.

	Seeks supervision/consultation and mentorship as necessary to enhance competence in research.

Competency B: ETHICAL AND LEGAL STANDARDS

B1. Goal: Patient Risk Management and Confidentiality

	Effectively evaluates, manages and documents patient risk in terms of immediate concerns such as suicidality, homicidality, and any other safety issues.

Specific Objectives:

	Assesses all risk situations fully prior to leaving work site for the day.
	Collaborates with patients in crisis to make short-term safety plans, and intensify treatment as needed.
	Takes appropriate actions to manage high risk situations (e.g. escorting patients to ER) immediately in a manner consistent with departmental/institutional policy.
	Follows up with patients, collaterals, and/or other health professionals appropriately.
	Documents all high risk situations and their management appropriately and promptly.
	Seeks and utilizes supervision/consultation appropriately.

B2. Goal: Knowledge of Ethics and the Law

	Demonstrates good knowledge and appropriate application of ethical principles and state law.

Specific Objectives:

	Identifies ethical and legal issues spontaneously and consistently, and addresses them proactively and thoughtfully.
	Uses good and reliable judgment about when supervision/consultation is needed.
	Is responsive to supervisory input and utilizes information appropriately.

Competency C: INDIVIDUAL AND CULTURAL DIVERSITY
C1. Goal: Sensitivity to Patient Diversity

	Exhibits sensitivity to the individual and cultural diversity of patients and commitment to providing culturally sensitive services.

Specific Objectives:

	Understands issues involved in working with patients of diverse backgrounds and characteristics.
	Acknowledges and respects differences that exist between self and patients in terms of race, ethnicity, culture, and other individual difference variables.
	Discusses individual difference variables with patients when appropriate.
	Recognizes when more information is needed regarding patient differences and seeks out information autonomously.
	Recognizes own limits to expertise and seeks supervision/consultation as necessary.
	Is able to work effectively with patients who have diverse backgrounds and characteristics.
	Demonstrates knowledge of health disparities particularly as it applies to relevant vulnerable populations.

C2: Goal: Awareness of Own Cultural and Ethnic Background

	Demonstrates awareness of own background and its impact on patients, and exhibits commitment to exploring these variables in relation to clinical practice.

Specific Objectives:

	Accurately monitors own responses to differences, and differentiates these from patient responses.
	Exhibits awareness of personal impact on patients different from self.
	Demonstrates willingness to be thoughtful about own cultural identity and other individual difference variables.
	Reliably seeks supervision/consultation as necessary and utilizes feedback.

Competency D: PROFESSIONAL VALUES, ATTITUDES, AND BEHAVIORS

D1: Goal: Professional Interpersonal Behavior

	Professional and appropriate interactions with treatment teams, peers, supervisors, and other professionals.

Specific Objectives:

	Has smooth working relationships with peers, supervisors, and other professionals.
	Handles differences openly, tactfully, and effectively.
	Participates actively and behaves professionally in staff meetings, seminars, lectures, case conferences, and other settings.

D2: Goal: Professional Responsibility

	Demonstrates responsibility for key patient care tasks which are completed promptly.

Specific Objectives:

	Maintains complete records of all patient contacts and pertinent information.
	Produces clear and concise progress notes with appropriate attention to detail.
	Completes all documentation in a timely manner.
	Takes initiative in ensuring that key tasks are accomplished.
	Ensures that records always include crucial information.

D3: Goal: Efficiency, Administrative Competency, and Time Management

	Demonstrates efficient and effective time management.

Specific Objectives:

	Efficiently completes tasks without prompting, deadlines or reminders.
	Utilizes time management skills regarding appointments, meetings, and leave.
	Follows departmental and institutional policy regarding scheduling, vacations, and other responsibilities.
	Independently assesses the larger task to be accomplished, breaks task down into smaller components, and develops timetable for completion.

	Prioritizes various tasks and deadlines efficiently and without need for supervisor input.
	Makes adjustments to priorities as demands evolve.

D4: Goal: Use of Reflective Practice, Self-Assessment, and Self-Care in Professional Development

	Engage in reflective practice conducted with self-assessment to further personal and professional development.

Specific Objectives:

	Appreciates and attends to own health behaviors and well-being and their potential impact on practice.
	Exhibits good awareness of personal and professional problems.
	Manages stressors to minimize impact on professional practice.
	Is open and non-defensive in soliciting and incorporating feedback and recommendations from supervisors and other professionals.
	Demonstrates positive coping strategies with personal and professional stressors and challenges.
	Actively seeks supervision/consultation and/or personal therapy to resolve relevant issues.
	Appreciates the importance of professional development and utilizes supervision/mentorship appropriately

Competency E: COMMUNICATION AND INTERPERSONAL SKILLS

E1: Goal: Effective Interpersonal Communication

	Demonstrates effective communication in multiple settings and roles.

Specific Objectives:

	Relates effectively and appropriately with patients, colleagues, supervisors, and other health professionals.
	Communicates clearly and appropriately in written and oral form with patients, colleagues, supervisors, and other health professionals.
	Demonstrates the ability to form alliances, deal with conflict, negotiate differences, and understand and maintain appropriate professional boundaries with patients, colleagues, supervisors, and other health professionals.

E2: Goal: Use of Supervision/Consultation

	Seeks supervision/consultation as necessary and uses it productively.

Specific Objectives:

	Actively seeks supervision/consultation when treating complex cases and in unfamiliar circumstances.
	Prepares for supervision in order to maximize usefulness of consultation.
	Communicates effectively with supervisor in order to obtain necessary support/information.
	Provides supervisor with necessary materials in preparation for supervisory sessions.

	Incorporates supervisory input into work.
	Is appropriately assertive and not overly deferential towards supervisor.
	Is not overly defensive, is willing to accept feedback and to be observed and evaluated.

Competency F: PSYCHOLOGICAL ASSESSMENT AND DIAGNOSIS

F1. Goal: Assessment and Diagnostic Skills

	Demonstrates a thorough working knowledge of psychological assessment, psychiatric diagnostic nomenclature, and DSM, ICD, and other (e.g., ICSD) classifications relevant to general outpatient and specialty clinical populations (e.g., addiction, weight management, sleep, oncology, etc).

Specific Objectives:

	Identifies and obtains necessary information, including information from sources other than the interview (previous records, collateral, information, etc.).
	Utilizes historical, interview and psychometric data to diagnose accurately.
	Integrates data from various sources into a coherent conceptualization of the patient using a biopsychosocial formulation.
	Selects appropriate diagnosis and is able to support diagnosis with data indicating that the diagnostic criteria have been met.
	Conducts evaluations and provides assessments grounded in evidence-based practice.
	Uses assessment data including objective measures (e.g., self-report measures, actigraphy, food and sleep logs, etc.) to develop an appropriate and realistic treatment plan collaboratively with patient and accounting for patient's developmental and cognitive level.
	Develops comprehensive and concise assessment reports in a timely manner.
	Communicates findings from assessments to patients, collaterals, and other health professionals as necessary.
	Identifies and responds appropriately to situations requiring immediate intervention (e.g. risk) and follows up as necessary.

F2. Goal: Consultation/Liaison Assessment and Interview Skills
(Complete for Adult/Pediatric C/L and other Multidisciplinary Medical Clinic Rotations)

	Ability to conduct a comprehensive interview, considering reason for consult request and adapt interview style for medically ill patients.

Specific Objectives:

	Understands the referral question and relevant medical information/diagnostic testing.
	Identifies and obtains necessary information, including information from sources other than the interview (previous records, chart, hospital staff, family, etc.)
	Adapts interview style for medically ill patients in a variety of settings (i.e. patients on ventilators, patients in isolation rooms, recognize stress and fatigue in patients) and prioritizes questions.
	Recognizes areas requiring more in depth inquiry based on reason for consult and history.
	Covers all areas in interview such as, history of present illness, past medical history, past psychiatric history, family history, social history, academic history, developmental history, and mental status.
	Interview and report is organized and flows to pertinent topics.
	Identifies and responds appropriately to situations requiring immediate intervention such as a sitter or hospitalization (e.g. risk).
	Is able to clearly provide feedback to the consulting team with case-conceptualization and recommendations.
	Identifies and obtains necessary information, including information from sources other than the interview (previous records, chart, hospital staff, family, etc.).
	Produces comprehensive documentation (e.g., consultation note, C/L report) that is clear, concise and with appropriate recommendations.

F3. Goal: Neuropsychological Assessment
(Complete for Adult and Pediatric Neuropsychology Rotations)

	Ability to promptly and efficiently administer and score commonly used neuropsychological tests

Specific Objectives:

	On time, prepared, and organized to begin the testing
	Appropriately selects tests, considering patient characteristics, including issues of diversity and referral questions
	Establishes appropriate relationship with patient before and during the assessment
	Administration is consistent with manual instructions and appropriately time sensitive
	Scoring follows guidelines and is accurate
	Appropriate selection of norms: awareness of limitations and strengths of norms
	Appropriate selection of norms: awareness of limitations and strengths of norms and their applicability to patient's age and developmental level
	Obtains data from collaterals and other providers as needed

F4. Goal: Neuropsychological Interview
(Complete for Adult and Pediatric Neuropsychology Rotations)

	Ability to conduct a comprehensive interview, considering referral question, neuropsychology principles, medical risk factors, and knowledge of neuroanatomy

Specific Objectives:

	Covers all areas: Presenting complaint, medical, psychiatric, developmental, social, academic, & work history, premorbid functioning, personality
	Interview is organized and flows to pertinent topics
	Recognizes areas requiring in depth inquiry based on presenting complaint, history, and medical aspects of case
	Maintains a good rapport with the patient and guardian(s)
	Solicits appropriate input from available collateral source(s)
	Conducts relevant neurobehavioral examination
	Identifies and responds appropriately to situations requiring immediate intervention (e.g. risk)

F5. Goal: Production of Neuropsychological Report
(Complete for Adult and Pediatric Neuropsychology Rotations)

	Ability to produce a comprehensive, insightful, and accurate report (level of sophistication, comprehensiveness, clarity of communication, conciseness, specific recommendations for clinical care, attention to detail)

Specific Objectives:

	Identifies, obtains, and integrates information from sources other than the interview (previous records, collateral information, diagnostics, NP history questionnaire, etc.)
	Basic writing is concise, organized, comprehensive, accurate, and relevant for referral source
	Information gleaned from interview is organized, reflecting relevant facts and knowledge (Background information section of report)
	Appropriate interpretation of neuropsychological test data based on relevant norms (Test Results section)
	Appropriate interpretation of personality and behavioral test data based on relevant norms (Test Results section)
	Demonstrates an understanding of related medical records and diagnostic test results (CT, MRI, etc)
	Integrates test data, with referral question, presenting complaint, diagnostics, and history into a coherent conceptualization of the client (Summary section)
	Selects appropriate diagnosis and is able to support diagnosis with data indicating that the diagnostic criteria have been met
	Recommendations are sensitive to referral question, conceptualization and patient specific variables and resources.

F6. Goal: Neuropsychological Knowledge
(Complete for Adult and Pediatric Neuropsychology Rotations)

	Sophisticated and comprehensive knowledge of (developmental) neuropsychological principles, and related areas pertinent to assessment and diagnosis (i.e., behavioral and emotion regulation; developmental psychopathology; interventions for home/work/school, etc.).

Specific Objectives:

	Comprehensive understanding of common neurological and neurodevelopmental diseases and conditions
	Comprehensive understanding of common medical conditions in childhood and adolescence that impact cognition and behavior
	Demonstrates a clear understanding of common neurocognitive profiles related to specific diseases and neurodevelopmental conditions
	Basic understanding of medication and side effects
	(Developmental) Neuroanatomy knowledge

Competency G: PSYCHOTHERAPEUTIC INTERVENTION

G1: Goal: Patient Rapport

	Consistently achieves good rapport with patients and collaterals.

Specific Objectives:

	Establishes and maintains a working relationship with most patients and collaterals.
	Is aware of relationship issues which may impact the course of treatment and manages these issues effectively.
	Recognizes the boundaries of the therapeutic relationship and the therapist's appropriate responsibilities.
	Reliably identifies potentially challenging patients and seeks supervision/consultation as necessary.

G2: Goal: Case Conceptualization and Treatment Planning

	Develops a useful biopsychosocial case conceptualization that draws on theoretical and research knowledge.

Specific Objectives:

	Formulates a good biopsychosocial case conceptualization within own preferred evidence-based theoretical orientation.
	Demonstrates flexibility and is able to appreciate insights from other evidence-based theoretical orientations.
	Has working knowledge of medical diagnoses relevant to patient's presentation and can integrate this information into case formulation and treatment planning.
	Collaborates with patient and other providers as necessary to develop and communicate appropriate and short-, intermediate- and long-term treatment goals.
	Monitors and revises plan appropriately in response to patient's progress.
	Identifies issues or long-range goals which are not related to current treatment environment and makes referrals appropriately

G3: Goal: Psychotherapy

	Interventions are well-timed, effective, and consistent with empirically supported treatments and/or other psychological literature.

Specific Objectives:

	Demonstrates flexibility and is able to apply differential methods of treatment.

	Selects and implements interventions appropriate to the patient(s).
	Conducts therapy in an organized and focused way, consistent with the treatment plan.
	Uses interventions and interpretations to facilitate patient acceptance and change.
	Integrates use of self-help groups, bibliotherapy and other referrals/resources with psychotherapeutic approach.
	Recognizes when patient needs more or less restrictive levels of care and manages transition effectively.
	Understands and uses own emotional reactions to the patient productively in the treatment.
	Presents appropriate interpretations to supervisor and/or patient as necessary.
	Seeks supervision/consultation as necessary for complex cases.

G4: Goal: Consultation/Liaison Interventions
(Complete for Adult/Pediatric C/L and other Multidisciplinary Medical Clinic Rotations)

	Demonstrates knowledge of the general principles of the individual and family aspects of the psychology of medical illness and is able to provide brief supportive psychotherapy at bedside

Specific Objectives:

	Demonstrates flexibility and is able to apply differential methods of treatment, especially when working at the bedside
	Understand the use of psychosocial treatments including brief psychotherapy, behavioral management techniques, family therapy, and psychoeducation.
	Selects and implements interventions appropriate to the patient(s) and family.
	Demonstrates motivation to increase knowledge and expand range of interventions through reading and consultation as necessary.

G5: Goal: Case Management

	Makes appropriate referrals to meet patient's needs, provides consultation as necessary, and collaborates with other professionals.

Specific Objectives:

	Recognizes the boundary of one's own limitations in treating particular patients and makes appropriate referrals as necessary.
	Obtains and provides referrals and follow up services to patients, collaterals, and other professionals as appropriate.
	Documents all case management activities.
	Seek supervision/consultation as necessary for complex cases.

G6: Goal: Group Therapy

	Functions effectively as group (co)facilitator.

Specific Objectives:

	Understands and maintains appropriate structure for group, according to group membership and purpose.
	Elicits participation and cooperation from all members.

	Recognizes group process and is able to use it to facilitate patient progress in group.
	Selects and implements interventions which facilitate group process and patient progress.
	Prepares independently for each session with little or no prompting.
	Confronts group problems appropriately and independently.
	Establishes and maintains productive working relationships with co-therapist(s).
	Seeks supervision/consultation and uses incorporates feedback into work.
	Manages group alone in the absence of co-therapist/supervisor with follow-up supervision later on.
	Completes documentation of group notes and follow-up contact in an appropriate and timely manner.

COMPETENCY H: SUPERVISION, EDUCATION, AND TRAINING

H1: Goal: Supervision Skills

	Demonstrates good knowledge and use of supervision theory, models, techniques, and skills.

Specific Objectives:

	Engages in independent efforts to learn about supervision theory, models, and effective practices in supervision (e.g., directed readings).
	Is knowledgeable about theories, models, and effective practices in supervision.
	Spontaneously and consistently applies supervision skills.
	Builds rapport, establishes working relationship, and is appreciated by supervisee.
	Provides useful direction, information, and feedback that is appropriate for supervisee's developmental level.
	Seeks supervision/consultation as necessary.

H2: Goal: Training/Teaching Skills

	Demonstrates training/teaching skills in a clinical-educator capacity.

Specific Objectives:

	Expresses interest in and seeks opportunities to develop teaching/training skills.
	Provides effective presentations in courses and didactic sequences.
	Facilitates audience engagement by adapting didactic information as necessary.
	Responds appropriately to audience's comments and questions.

COMPETENCY I: CONSULTATION AND INTERPROFESSIONAL/INTERDISCIPLINARY COLLABORATION

I1: Goal: Consultation Skills

	Provides consultation as necessary and collaborates with other professionals, including functioning as part of a multidisciplinary team.

Specific Objectives:

	Appreciates the importance of interprofessional practice including values, ethics, and roles of different disciplines.

	Provides psychoeducation to members of other teams and assists them in managing patients with comorbid psychiatric diagnoses.
	Functions effectively and cooperatively as a member of a team working with professionals from other disciplines.
	Engages in appropriate and effective interactions with the consultees.
	Utilizes input from other members of the team in formulating findings and recommendations.
	Demonstrates awareness of the needs of the consulting team, especially when making recommendations and arranging for follow-up.
	Maintains consistent communication with the other members of the team (e.g., residents, fellows, attendings).
	Utilizes the electronic medical record to communicate appropriately as necessary.
	Seeks supervision/consultation as necessary for complex cases.
	Liaisons effectively with colleagues in other fields.

SUMMARY OF STRENGTHS:

AREAS FOR ADDITIONAL DEVELOPMENT OR REMEDIATION, INCLUDING RECOMMENDATIONS:

TRAINEE COMMENTS:

CONCLUSIONS

Remedial Work Instructions: In the rare situation when it is recognized that a trainee needs remedial work, a competency assessment form should be filled out **immediately**, prior to any deadline date for evaluation, and shared with the trainee and the Director of Psychology Training. In order to allow the trainee to gain competency and meet passing criteria for the training activity, these areas must be addressed proactively and a remedial plan needs to be devised and implemented promptly

Goal for Practicum Evaluations: All competency areas will be rated at a level of **2** or higher. No competency areas will be rated as **1**.

Goal for Internship Evaluations done prior to 12 Months: All competency areas will be rated at a level of competence of **I** or higher. No competency areas will be rated as **1** or **2.**

Goal for Internship Evaluations done at 12 Months: At least 80% of competency areas will be rated at level of competence of **4** or higher. No competency areas will be rated as **1** or **2**. Note: Exceptions would be specialty rotations that would take a more intensive course of study to achieve this level of competency and the supervisor, training director and trainee agree that a level of **3** is appropriate for that particular experience (e.g. C/L rotation for a trainee who has never worked with seriously medically ill patients)

Goal for Postdoctoral Training Evaluations done prior to 12 Months: 80% of competency areas will be rated at a level of competence of **4** or higher. No competency areas will be rated as **1** or **2**

Goal for Postdoctoral Training Evaluations done at 12 Months: At least 80% of competency areas will be rated at level of competence of **5** or higher. No competency areas will be rated as lower than **4**.

Check one:

☐ The trainee HAS successfully completed the above goal. We have reviewed this evaluation together.

☐ The trainee HAS NOT successfully completed the above goal. We have made a joint written remedial plan as attached, with specific dates indicated for completion. Once completed, the rotation/training experience will be re-evaluated using another evaluation form. We have reviewed this evaluation together.

Supervisor _____ Date _____

I have received a full explanation of this evaluation. I understand that my signature does not necessarily indicate my agreement.

Trainee _____ Date _____