

8-1-2012

Relationship Competence: Can Trainee Interpersonal Skills Be Measured Reliably and Do They Predict Clinical Effectiveness?

Jacqueline Camp
University of Nevada, Las Vegas

Follow this and additional works at: <https://digitalscholarship.unlv.edu/thesesdissertations>



Part of the [Clinical Psychology Commons](#), and the [Counseling Psychology Commons](#)

Repository Citation

Camp, Jacqueline, "Relationship Competence: Can Trainee Interpersonal Skills Be Measured Reliably and Do They Predict Clinical Effectiveness?" (2012). *UNLV Theses, Dissertations, Professional Papers, and Capstones*. 1657.
<http://dx.doi.org/10.34917/4332638>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by Digital Scholarship@UNLV with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you need to obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/or on the work itself.

This Dissertation has been accepted for inclusion in UNLV Theses, Dissertations, Professional Papers, and Capstones by an authorized administrator of Digital Scholarship@UNLV. For more information, please contact digitalscholarship@unlv.edu.

RELATIONSHIP COMPETENCE: CAN TRAINEE INTERPERSONAL
SKILLS BE MEASURED RELIABLY AND DO THEY
PREDICT CLINICAL EFFECTIVENESS?

by

Jacqueline Camp

Bachelor of Arts in Psychology
University of Nevada, Las Vegas
2004

Master of Arts in Psychology
University of Nevada, Las Vegas
2007

A dissertation submitted in partial fulfillment of
the requirements for the

Doctor of Philosophy Degree in Psychology
Department of Psychology
College of Liberal Arts

Graduate College
The University of Nevada, Las Vegas
August 2012



THE GRADUATE COLLEGE

We recommend the dissertation prepared under our supervision by

Jacqueline Camp

entitled

Relationship Competence: Can Trainee Interpersonal Skills Be Measured Reliably and Do They Predict Clinical Effectiveness?

be accepted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Psychology

Department of Psychology

Christopher Heavey, Committee Chair

Michelle Carro, Committee Member

Marta Meana, Committee Member

Kimberly Barchard, Committee Member

Stephen Fife, Graduate College Representative

Thomas Piechota, Ph. D., Interim Vice President for Research and Graduate Studies
and Dean of the Graduate College

August 2012

ABSTRACT

Relationship Competence: Can Trainee Interpersonal Skills be Measured

Reliably and Do they Predict Clinical Effectiveness?

by

Jacqueline P. Camp

Dr. Christopher Heavey, Examination Committee Co-Chair
Associate Professor of Psychology
University of Nevada, Las Vegas

Traditional evaluation and assessment procedures in professional psychology programs have long been criticized for inadequately attending to the set of interpersonal skills that are important to professional functioning in the field of psychology. The present study was exploratory and focused on examining the reliability and validity of an evaluation tool designed to capture a set of interpersonal skills that are clinically relevant and grounded in the empirical literature on psychotherapy outcome. Toward this end, the Facilitative Interpersonal Skills (FIS) task (Anderson, Patterson, & Weiss, 2006) was administered to a sample of trainees ($n = 19$) enrolled in a clinical psychology doctoral program and a marriage and family therapy master's program. The FIS task is a performance based evaluation method that attempts to measure interpersonal behavior samples taken in response to videotaped vignettes simulating challenging therapeutic situations. Trainee interpersonal responses to the task were evaluated and rated on the basis of the FIS scale by four independent raters. Other measures that could potentially be used to evaluate trainee performance or relevant skills were gathered, including measures of academic ability/performance, quantity of experience, self-reported interpersonal skills, and client outcome. Consistent with previous research, results

indicated that ratings of trainee performance on the FIS task could be made reliably.

With respect to validity, better performance on the FIS task was associated with more years in training, particularly for clinical psychology trainees; and unrelated to measures of academic ability. Results involving measures of client outcome were deemed inconclusive due to very small sample size, missing data, and other concerns. Findings are discussed in terms of implications for improving current training evaluation and assessment practices in professional psychology training programs.

ACKNOWLEDGEMENTS

I am deeply grateful to my advisors and committee chairs, Drs. Chris Heavey and Michelle Carro, for their insight and guidance throughout this project and across my graduate career. You have inspired and helped me more times than I can count over the course of a challenging eight years. Thank you for your genuine investment, patience, and steadfast belief in me. I am also greatly appreciative for the significant and insightful contributions of my committee members, Drs. Marta Meana, Kim Barchard, and Stephen Fife. This project would not have been possible without Brittany Edwards, who spent long, arduous hours working on data collection. Thank you for your reliability, stability, attention to detail, and dedication to completing this project. I also owe a huge thanks to Brittney Rufkahr, for her assistance with data collection and to my coding team, Sarah Akhter, Shamell Brandon, Ladonna Hayden, and Erin Snyder for their amazing coding talents. Finally, my deepest and sincerest gratitude and love to: my husband and best friend, Seth, for his love, unshakable stability, and constant support and safety; to my best friend and soul-sister, Paula for helping me to see beauty and light in the dark, as well as the forest for the trees; to my dearest friends, Sarah, Caitlin, Sharon, Jelly Belly, and Arva for their love, laughs, and smiles; to my Ozzy, Sam, and Lucky dogs who have made me smile when I didn't think it was possible and who have taught me the meaning of love; and to my family (both blood and commune) for their constant support and patience. This dissertation is dedicated to my mother, Barbara Ann Camp, and my father, Jerome (Jerry) James Camp, who showed me how to live. May their fierce love and unfaltering belief in me continue to help me to live a life which would make them proud. Their memories will live on in my heart and in every moment of my life.

TABLE OF CONTENTS

ABSTRACT	iii
ACKNOWLEDGEMENTS	v
LIST OF TABLES	vii
CHAPTER 1 INTRODUCTION	1
CHAPTER 2 LITERATURE REVIEW	6
Early Evaluation of Professional Psychology Trainees	6
Empirical Directions for Identifying Promising Professionals	19
Therapist Interpersonal Skills and Clinical Effectiveness	33
Empirical Directions for Evaluating Interpersonal Skills	58
The Present Study	67
CHAPTER 3 METHODOLOGY	71
Participants	71
Instruments	74
Procedure	83
CHAPTER 4 RESULTS	88
Reliability of Ratings of Interpersonal Performance	88
Interpersonal Performance and Stable Trainee Characteristics	91
Interpersonal Performance and Traditional Evaluation Methods	92
Interpersonal Performance and Clinical Effectiveness	95
CHAPTER 5 DISCUSSION	101
Primary Findings	101
Limitations	113
Implications for Practice	115
Implications for Research	118
APPENDICES	120
Appendix A: Demographic Questionnaire	120
Appendix B: Facilitative Interpersonal Skills Task Sample Vignettes	123
Appendix C: FIS-O Scale Correlation Matrix	125
Appendix D: Therapist Interpersonal Skills Self-Report (TIS-SR) Inventory	126
Appendix E: Correlations: FIS-O and TIS-SR	129
REFERENCES	130
VITA	174

LIST OF TABLES

Table 1	FIS Performance Condensed Mean Total and Item Scores	78
Table 2	Inter-rater Reliability: FIS Ratings	88
Table 3	Inter-rater Reliability: FIS Ratings for 4-Raters, 3-Rater Combinations, and 2-Rater Pairs	89
Table 4	Inter-rater Reliability: FIS Ratings, 3-Rater Combinations	90
Table 5	Inter-rater Reliability: FIS Ratings, 2-Rater Pairs	90
Table 6	Academic Performance/Aptitude and Interpersonal Performance	93
Table 7	Year in Training and Interpersonal Performance.....	94
Table 8	Therapy Client OQ-45 Change Scores and Percentage of Caseload Missing.....	96
Table 9	Interpersonal Performance and Average Caseload Change.....	97
Table 10	Therapist Interpersonal Skills Self Report and Average Caseload Change	98
Table 11	Therapy Specific and Traditional Therapist Variables and Average Caseload Change.....	99
Table 12	Academic Performance/Aptitude and Average Caseload Change....	100

CHAPTER 1

INTRODUCTION

Training programs are gatekeepers of the profession of psychology and are ethically obligated to assure excellence in the individuals who are allowed passage into the profession, as well as to protect the public from those individuals who are not appropriate for professional functioning (Behnke, 2005). Reliable and meaningful evaluation tools represent the foundation on which adequate gatekeeping rests, and in which the quality of the profession is partially determined (APA, 2006a; Kaslow, Rubin, Forrest, Elman, VanHorne, Jacobs et al., 2007; Rubin, Bebeau, Leigh, Lichtenberg, Nelson, & Portnoy, 2007). Despite the importance of reliable and valid evaluation tools, questions have arisen regarding the utility of the traditional evaluation methods commonly used in professional training programs (APA, 2006a; Kaslow et al., 2007; Nelson, 2007; Rubin et al., 2007).

Overall, recent critiques suggest that standard evaluation methods utilized in graduate training programs overemphasize academic abilities, but underemphasize those personal abilities (e.g., interpersonal, intrapersonal, emotional skills) that decades of research suggest are “inextricably intertwined” in the roles that professional psychologists fill (Norcross, 2002, p. 4; see also APA, 2006a; Assessment of Competency Benchmarks Work Group, 2007). The overemphasis on academic abilities is apparent during the admission process, in which selection criteria are heavily biased toward selecting students who are intelligent and academically proficient (Peterson, 2003). In contrast, selection criteria do not adequately attend to students’ capacity to be talented or even competent clinicians. Biased attention toward academic abilities at admission continues throughout

training, perpetuated by reliance on coursework performance and quantity of clinical experience, despite empirical evidence that directly contradicts the notion that passage of courses and acquisition of clinical hours are adequate benchmarks for professional competence (APA, 2006a; Hatcher & Lassiter, 2007; Kaslow et al., 2007; McHolland, Peterson, & Brown, 1987; Rubin et al., 2007).

The relative neglect of trainee personal characteristics and abilities seems partially related to the challenge inherent in defining and adequately capturing the personal abilities relevant to successful clinical performance (APA, 2006a; Lichtenberg, Portnoy, Bebeau, Leigh, Nelson, Rubin et al., 2007). As a result, empirically supported tools for assessing relevant trainee personal skills largely do not exist. Despite these challenges, decades of industrial organizational research on performance-based appraisal methods (Schmidt & Hunter, 1998) suggests that early training evaluation methods could be improved by attending to those skills and abilities that lead to enhanced performance in the complex professional roles that trainees will be expected to fill (Peterson, 2003). This suggests that it is possible (and potentially valuable) to identify, delineate, and evaluate those characteristics (at selection and throughout training) that might distinguish those trainees who are interpersonally capable from those who will fall short of what is required for proficient work, and in fact, whose skill deficits might cause harm to others.

Along these lines, a great deal of conceptual and empirical attention has been focused on the professional role of psychotherapist and is directly relevant to determining which personal skills and abilities might separate those trainees who are talented and more likely to be clinically effective, from those who are less so. In essence, the psychotherapy research literature suggests that the personal, working relationship

between therapist and client is an important determinant of psychotherapy outcomes (Norcross, 2002; Orlinsky, Ronnestad, & Willutzki, 2004). More importantly, this research indicates that the therapist makes interpersonal contributions to the strength and quality of the therapy relationship, contributions which are interwoven with the effectiveness of specific psychotherapy interventions and treatments (Beutler, Malik, Alimohamed, Harwood, Talebi, Noble et al., 2004; Castonguay, Boswell, Constantino, Goldfried, & Hill, 2010; Norcross, 2002).

Differences between therapists who are effective and those who are less effective have often been unrelated or weakly related to years of experience (e.g., accumulation of practicum hours), approach to treatment, or technical skills (e.g., learned in courses or manuals). Instead, variation among therapists seems attributable to the therapists' ability to build a high quality, helpful relationship with another person (Asay & Lambert, 2002; Lambert, 1989; Lambert & Okiishi, 1997; Rogers, 1957; Strupp, 1995). Overall, decades of research taken together suggests that individuals who are empathic, warm, and able to successfully withstand and negotiate conflict in an open, understanding, confident, expert, focused manner are more successful in professional interactions, and are better equipped to effectively facilitate the interpersonal process of psychotherapy (Norcross, 2002). In short, it may be that trainees who possess these interpersonal abilities at the outset of training are most suitable for the profession, and with appropriate training and evaluation, are likely to become effective therapists (Hatcher & Lassiter, 2007).

Recent criticisms of training evaluation practices combined with clear empirical directions about the interpersonal skills that are fundamental to clinical competence suggest that it is time to examine new evaluation methods that are more relevant to

predicting professional performance (Kaslow et al., 2007; Lichtenberg et al., 2007; Roberts, Borden, Christiansen, & Lopez, 2005; Rubin et al., 2007). A paucity of research has addressed the need for reliable, valid, and relevant evaluation tools intended for use during graduate training (APA, 2006a). Further, the few evaluation tools that have been proposed for use in professional training lack empirical support with regard to reliability and validity, or are long outdated. Therefore, the goal of this dissertation is to evaluate the psychometric properties and predictive utility of an evaluation method that attends to fundamentally important interpersonal skills and may be useful for adequately measuring the relational ability of trainees in professional psychology training programs.

The importance of evaluating trainee interpersonal abilities is not simply an issue of academic debate. As noted earlier, evaluation in professional psychology programs partially shapes the future of the discipline and as such, has implications for the effectiveness of training, as well as for the public that professional psychologists serve (Kaslow et al., 2007). Relevant, standardized, and psychometrically strong evaluation tools could enhance recognition of trainee problems with interpersonal competence and trainers' ability to adequately address problems (through remediation or removal) without fear of legal ramifications. In other words, trainers would have a means to operationalize the interpersonal skills that are, at times, lacking in trainees, but often subjectively evaluated on a "gut-level" which is difficult to justify in a court of law (Forrest, Elman, Gizara, & Vacha-Haase, 1999). Moreover, identifying unsuitable trainees at the outset of training would better ensure that trainers are teaching those who are most likely to learn what it is they have to teach, rather than those who come without the necessary baseline interpersonal skill set (i.e., garbage in, garbage out). Finally, the interpersonal abilities of

professional therapists are important determinants of clinical outcomes, such that interpersonally unskilled therapists may be less able to facilitate client change, and interpersonal skill deficits may even be related to client deterioration (Lambert, 1989).

This literature review is presented in four sections. The first section provides an overview of current training evaluation methods and procedures, particularly those at admission, and reviews research suggesting that current early evaluation procedures are inadequate for assessing trainee interpersonal skills. The second section reviews research regarding the influential role of therapist relational ability in shaping psychotherapy outcomes. This research challenges the current overemphasis in training programs on academic capacities and suggests that interpersonal abilities partially determine professional effectiveness, despite evaluation methods that fail to consider these skills. The third section provides an empirical review of the interpersonal skills that have been most consistently linked with professional outcomes and are essential to assess early on in the training process. The fourth section presents a promising approach for effectively measuring trainee interpersonal skills and offers a refined view of how these interpersonal skills might be adequately captured early on in the training process.

CHAPTER 2

LITERATURE REVIEW

Early Evaluation of Professional Psychology Trainees

Admission to a professional psychology training program is the first of several gateways through which an individual must successfully pass in order to enter into the profession of psychology. Ideally, admission serves to “exclude would-be psychologists who do not possess required intellectual and personal qualities” and to “recruit candidates who offer promise of outstanding performance” (Peterson, 2003, p. 797). Evaluation throughout training serves to assure that those trainees who have been selected are actually those who offer promise. However, typical selection criteria and evaluation procedures in professional psychology programs have been criticized for not adequately attending to all of the essential skills and qualities that candidates need to become successful professional psychologists.

Selection of Intelligent Students

Admission criteria for professional psychology graduate training programs are largely ignored in formal training guidelines (APA, 2008). Programs are only instructed to show that their graduate students “by interest, aptitude, and prior achievement are of quality appropriate for the program’s goals and objectives” (APA, 2008, p. 12). Given the large number of applications received and the relatively small portion of students that are selected for graduate study (10% in accredited clinical PhD and 40% in APA accredited PsyD programs; Norcross, Kohout, & Wicherski, 2005) programs have the latitude to be highly selective at admissions. Applicants typically provide three letters of recommendation, Graduate Record Examination (GRE) scores, a personal statement that

outlines previous accomplishments and future goals, an academic vita, and undergraduate grade point average (GPA; Fauber, 2006). Standard practice is to screen students from the larger pool of applicants (90% of PhD and PsyD programs screen applicants based on minimum GRE scores; Mayne, Norcross, & Sayette, 1994) and invite a smaller pool of applicants for on-site interviews with program faculty (Fauber, 2006).

Graduate departments differ in the importance they place on specific admission variables, but aggregate survey data based on program self-report suggest that letters of recommendation, the personal statement, GPA, and the interview are the most important, followed by research experience, GRE scores, and clinical service, and then work experience (Norcross, Kohout, & Wicherski, 2005). Somewhat inconsistent with the program-rated importance of GRE scores in survey data (Norcross, Kohout, & Wicherski, 2005) applicants are typically screened on the basis of the GRE and GPA (Fauber, 2006; Mayne, Norcross, & Sayette, 1994). The statement of purpose, letters of recommendation and prior experiences are relied upon to gauge “fit” with the graduate program, whereas variables such as the GRE and GPA are intended to capture aptitude for academic performance (Linn, 1990).

Decades of research suggests that intelligence (or general mental ability) is the strongest and most consistent predictor of future performance and capacity for learning in complex occupations (Schmidt & Hunter, 1998). Cognitive abilities, such as quantitative and verbal reasoning, academic abilities, such as writing skill, and subject specific aptitude taken together are inferred from scores on the GRE (Kuncel & Hezlett, 2007). Scores on the GRE are empirically related to admissions committee decisions (Ingram & Zurawski, 1981) and several outcomes, including first year grades (Hackman, Wiggins,

& Bass, 1970; Hirschberg & Itkin, 1978; Kuncel, Hezlett, & Ones, 2001; Sternberg & Williams, 1993), performance on comprehensive examinations (Dollinger, 1989; Kirnan & Geisinger, 1981) and favorable faculty ratings (Dollinger, 1989; Kuncel, Hezlett, & Ones, 2001).

Though much controversy has surrounded the use of indices of academic ability in predicting graduate student success in professional psychology (largely due to restriction of range; Huitema & Stein, 1993; Dawes, 1975), the GRE is arguably the most empirically supported variable utilized in the professional psychology admissions process. At minimum, GRE scores, and to a lesser extent GPA, provide an indication of whether prospective students have the aptitude to meet the academic demands of graduate study (Sternberg & Williams, 1993; Willingham, 1974). However, academic variables *alone* do not adequately capture whether students selected for graduate study are well-suited (or at minimum *capable*) to meet the demands of clinical training, or to achieve clinical competence (Bergin & Jasper, 1969; Bergin & Solomon, 1970; Constanzo & Philpott, 1986; Hosford, Johnson, & Atkinson, 1984; Hurst & Shatkin, 1974; Ingram, 1983; Kelly & Fiske, 1951; Littlepage, Bragg, & Rust, 1978; Loo, 1979; Omizo & Michael, 1979; Sternberg & Williams, 1993).

Unfortunately, the emphasis on intelligence and academic aptitude at admission appears to continue throughout training in professional psychology (APA, 2006a; Beutler, 1995; Bickman, 1999; Kaslow et al., 2007; Roberts et al., 2005). Trainers have a wide variety of evaluation methods to ensure that trainees are academically successful, including course grades, comprehensive exams, and research requirements (APA, 2006a). These evaluation methods ensure that trainees have the intelligence to acquire the body of

psychological knowledge necessary to be competent and master the complex professional skills needed to function effectively as a psychologist. However, achievement of high course grades and successful completion of research requirements (similar to GRE scores) do not sufficiently capture the full skill set necessary for adequate professional functioning (APA, 2006a; Beutler, 1995; Bickman, 1999; Kaslow et al., 2007; Ladany, 2007; McHolland, Peterson, & Brown, 1987).

Recommendations for Selecting All Around Promising Professionals

Although not all professional psychology careers involve clinical work, there is no career in applied psychology that does not involve working with others, and at least a large majority of applied psychologists, licensed or not, have an impact on the profession through service, training, applied research, or mentorship and teaching activities. Indeed, psychologists have emphasized the importance of selecting applicants with the appropriate personal characteristics for decades. Over 60 years ago, the committee of individuals who shaped training in clinical psychology emphasized that “the ability to carry out effectively the combination of functions called for depends upon the clinical psychologist’s being the right kind of person” who has a wide variety of characteristics, including “superior intellectual ability and judgment,” “interest in persons...a regard for the integrity of other persons,” and “ability to establish warm and effective relationships with others” (APA Committee on Training in Clinical Psychology, 1947, pp. 540-541).

Over 20 years later, psychologists were still asserting that individual students be “concerned, compassionate, intelligent, and sensitive” prior to selection because “training may mature and refine the experience of concern and empathy, but it cannot supply what does not exist in the first place” (Sakinofsky, 1979, p. 195; see also APA Committee on

Training in Clinical Psychology, 1947; Castonguay et al., 2010; Elman, Illfelder-Kaye, Robiner, 2005; Hatcher & Lassiter, 2007; Johnson & Campbell, 2002, 2004; Korman, 1974; NCSPP, 2007; O'Donovan & Dyck, 2001; Peterson, 2003; Spruill, Rozensky, Stigall, Vasquez, Bingham, & Olvey, 2004; Stricker & Callan, 1987). Modern psychologists continue to express concern about the need to select students who will be “outstanding” performers or, at minimum, competent professionals:

...decisions made regarding admittance into psychotherapy training programs are based on criteria that have little relation to... the potential for psychotherapy competence. Once in graduate school, it is rather difficult to gate-keep or even reroute students who may be deemed poor therapists (in large part because of variability in competence among the faculty and supervisors). It should not surprise us, then, that a decent percentage of students who graduate are not well equipped to be reasonably good therapists (Ladany, 2007, p. 395).

Overall, selection of students without proper attention to the essential personal characteristics for success in professional roles guarantees that at least a portion of those who graduate will be ineffective or minimally effective, and risks missing those who will “offer promise of outstanding performance” (Peterson, 2003, p. 797).

The problems with selection criteria are perpetuated throughout training with evaluation methods that do not adequately attend to trainee personal qualities. Indeed, there is evidence regarding the prevalence of trainees who exhibit problems with professional competence (Biaggio, Gasparikova-Kransnec, & Bauer, 1983; Shen-Miller et al., 2011). For instance, Forrest and colleagues (1999) provided estimates for the prevalence of competence problems on the basis of available literature and indicated that

65-77% of programs had students with “clinical deficiency” problems, 25-27% with ethical or unprofessional behavior problems, and 42-70% with interpersonal problems (Forrest et al., 1999). A more recent sample of training directors (n = 103) provided similar estimates (e.g., 65% inadequate clinical skills, 52% deficient interpersonal skills) and half had dismissed at least one trainee (most frequently because of problems with clinical skills) over the course of three years (Vacha-Haase, Davenport, & Kerewsky, 2004).

Once students are accepted into graduate programs, it is challenging to dismiss them, particularly on the basis of difficult to define problems, such as interpersonal qualities or clinical deficits (Forrest et al., 1999; Vacha-Haase, Davenport, & Kerewsky, 2004). As gatekeepers of the profession of psychology, training programs are ethically obligated to ensure that trainees can adequately perform relevant professional tasks. According to Behnke (2005), “authority rests with the gatekeeper to apply the criteria and so to allow, or not allow, passage” (p. 90). Training programs in psychology (and related helping fields) are ethically obligated to systematically evaluate trainees with respect to those skills that determine adequate functioning as early in the training process as possible (i.e., at admissions) and throughout training thereafter (from the earliest practicum training assignment to the start of internship).

Evaluating Clinical Competence at Admission and Beyond

Students admitted to professional programs are suited to succeed academically, and evaluation methods throughout training likely ensure that they do. In contrast, other relevant student characteristics, particularly those personal qualities that determine a candidates’ ability to relate well with others and behave professionally, have been

difficult to define and assess during early evaluation. Despite the notion that these qualities are more relevant to professional functioning or competence, current evaluation criteria utilized in training programs do not adequately attend to these abilities. Nonetheless, professional psychologists in any role should ideally possess certain personal qualities that if lacking, have the potential to reflect poorly on the discipline as a whole (Johnson & Campbell, 2002; 2004).

Johnson and Campbell (2002) refer to these qualities as “character” and “fitness” requirements for professional psychologists. They asserted that “essential character requirements” for professional psychologists include integrity, prudence, and caring (Johnson & Campbell, 2002, p. 406). The term “fitness” includes personality adjustment, emotional stability, and a lack of substance abuse (Johnson & Campbell, p. 406). The absence of these qualities in students is likely to lead to problems with professional competence (Johnson & Campbell, 2002) and arguably should be given serious consideration when excluding or removing students as part of the gatekeeping function of training.

Similarly, Hatcher and Lassiter (2007) provided a comprehensive outline of the skills and abilities that students ideally obtain during practicum training and highlighted “the need at the outset of professional training for evidence of various personality characteristics and intellectual and personal skills...” (p. 51). These qualities include: interpersonal skills (e.g., ability to listen, be empathic, respectful), cognitive skills (e.g., problem-solving, critical thinking), affective skills (e.g., ability to tolerate affect, conflict, and ambiguity), attitudes (e.g., desire to help, openness, honesty), expressive skills (e.g., communication ability), reflective skills (e.g., self-awareness), and personal skills (e.g.,

hygiene, appropriate dress). Presumably, identifying students with these abilities prior to training would result in the selection of individuals who would excel in meeting clinical training demands, and would later be the most successful as professional psychologists.

Generally, trainers recognize the importance of interpersonal abilities (Hatcher & Lassiter, 2007; Johnson & Campbell, 2002). For instance, in a recent survey of clinical training directors (n = 97), participants indicated that all of Johnson and Campbell's (2002) "essential character and fitness requirements" were important to assess at admission and throughout training, with the exception of substance use (Johnson & Campbell, 2004). Training directors most often reported using letters of recommendation, interviews, and personal statements at admissions to evaluate these interpersonal variables. They reported using clinical performance (presumably on the basis of clinical supervisor ratings), "personal behavior", advisor evaluations, faculty relations, and academic performance throughout training (Johnson & Campbell, 2004). Whereas predictors of academic success have demonstrated validity and reliability, and are routinely utilized in admissions and throughout training, there is limited evidence to suggest that current evaluation methods adequately capture potential for clinical ability.

Letters of Recommendation. Although rated as being the most often used by clinical training directors to screen for possible impairment in personal functioning (Johnson & Campbell, 2004) and ranked as the most important tool in selection (Norcross, Kohout, & Wicherski, 2005), letters of recommendation are invalid and unreliable indicators of the potential for success or for screening out applicants who are likely to become impaired (Aamodt, Bryan, & Whitcomb, 1993; Grote, Robiner & Haught, 2001; Hunter & Schmidt, 1998; Miller & Van Rycroek, 1988; Nicklin & Roch, 2009). Letters of

recommendation have been described as “wildly inflated” (Johnson & Campbell, 2004, p. 409; Miller & VanRybroek, 1988). Though letter-writers reported being candid about applicant weaknesses, this is inconsistent with letter-readers indication that letters rarely include negative information about applicants (Grote, Robiner, & Haught, 2001).

Letter inflation suggests that either all the individuals who apply for graduate school in professional psychology have the potential to be successful professionals, or that letters of recommendation do not differentiate between applicants. Consistent with the latter explanation, two studies on the relationship between letters of recommendation and faculty ratings of clinical ability/interpersonal skills in professional psychology programs indicated that letters were unrelated to clinical competency ratings (Federici & Schuerger, 1974; Piercy, Dickey, Case, Sprenkle, Beer, Nelson, & McCollum, 1995). Daehnert and Carter (1987) found evidence of a modest correlation between “dynamic and personal” letters of recommendation and supervisor ratings of intern motivation as well as practicum student responsibility and knowledge; but no evidence of a relationship between letters of recommendation and clinical skills. Taken together, extant research does not provide much support for the validity of letters of recommendation in capturing the personal skills required for the development of clinical competence.

The Personal Statement. The personal statement has also been questioned with regard to validity and utility for identifying candidates with strong personal skills (GlenMaye & Oaks, 2000; Powers & Fowles, 1997). Personal statements are written by applicants who strongly desire admittance and may present themselves in an overly positive manner. Furthermore, personal statements may not even be relevant to evaluating prerequisite baseline personal skills. For instance, one study suggested that

successful applicants were those who emphasized their commitment to research in their personal statements (Brown, 2004). Although this study is reflective of one research-oriented graduate program, this evidence is consistent with tacit knowledge among PhD applicants who are often advised to emphasize their interest in research (Fauber, 2006) and de-emphasize clinical interests. Only two empirical studies regarding the personal statement were located, and suggested that personal statements were unrelated to first year practicum performance (GlenMaye & Oakes, 2000) or clinical ratings made by faculty (Piercy et al., 1995).

Interview Day. Faculty-conducted interviews are also included as one of the most important criteria for providing information regarding applicant potential clinical abilities (Johnson & Campbell, 2004; Norcross, Kohout, & Wicherski, 2005). Typically, interpersonal and communication skills are evaluated on the basis of a brief, unstructured interview (Fauber, 2006) despite evidence that subjective clinical judgments are unreliable and invalid compared with actuarial prediction (Dawes, 1994; Kuncel, Hezlett, & Ones, 2001; Schmidt & Hunter, 1998). Fauber (2006) justified the use of unstructured admissions interviews to screen for impairment by using Meehl's (1957) observation that clinical judgment "works" in the case of unusual or extreme circumstances. In other words, clinical faculty members are able to make correct judgments when candidates have glaring social skill impairments (Fauber, 2006). Thus, interview day screening may serve to eliminate those applicants with gross impairment. Consistent with this, two studies have suggested that interview ratings were predictive of later faculty ratings of trainee interpersonal skills (Broadhurst, 1976; Federici & Schuerger, 1974).

Nonetheless, judgments made on the basis of brief interviews are likely biased by candidates “putting their best foot forward” and it is unclear how judgments of clinical ability or interpersonal skill are defined across faculty members (Nevid & Gildea, 1984). Consistent with this notion, several studies have failed to provide support for the predictive validity of interviews. In three studies, interviews were unrelated to later faculty ratings of clinical skills or professional performance (King, Beehr, & King, 1986; Piercy et al., 1995; Rikard & Clements, 1986). Taken together, these findings provide, at best, mixed support for the utility of the interview for predicting clinical skills.

Continuing the Troubling Trend. Inadequate attention to trainee interpersonal skills at admission continues throughout training, perpetuated by the “germ theory” and “practice makes perfect” myths of professional psychology training (Beutler, 1995, 1997; Bickman, 1999). Specifically, the “germ theory myth” refers to the reliance on coursework performance (e.g., “exposure to coursework leads to catching the skill bug”; Ladany, 2007, p. 392) to evaluate clinical competence. Similarly, the “practice makes perfect myth” refers to the tendency to infer clinical capability based on the quantity of clinical experience (e.g., the accumulation of practicum hours) without attention to the quality of training (Bickman, 1999; see also Beutler, 1997; Hatcher & Lassiter, 2007; Kaslow, Pate, & Thorn, 2005; Ko & Rodolfa, 2005; Lewis, Hatcher, & Pate, 2005; Rodolfa, Owen, & Clark, 2007).

These evaluation methods may be unsuitable for assessing clinical abilities, particularly given empirical evidence that directly contradicts these approaches (Atkins & Christensen, 2001; Buser, 2008; O’Donovan & Dyck, 2001; Ronnestad & Ladany, 2006; Stein & Lambert, 1995) and a paucity of evidence to justify that passage of courses and

acquisition of clinical hours are adequate benchmarks for achieving competence (APA, 2006a). Several empirical studies have failed to show that clinical experience significantly improves performance (for reviews see Beutler, 1995; Beutler et al., 2004; Bickman, 1999). For example, in a meta-analytic review of 36 studies, there was only a small positive relationship between experience and clinical outcomes ($ES = .20 - .30$; Stein & Lambert, 1995). Furthermore, preliminary evidence suggests that supervisor ratings, the primary means for evaluating clinical performance in professional psychology programs, are often inflated (overly positive) and unreliable (Borders & Fong, 1991; Ellis & Ladany, 1997; Gonsalvez, & Freestone, 2007).

Taken together, evidence suggests that interpersonal, character, and fitness qualities are evaluated unsystematically and subjectively during early evaluation in professional psychology programs. Furthermore, these subjective methods do not seem to predict important training outcomes, such as clinical competence or performance. Subjective methods are utilized partially because measures suitable for use in early evaluation do not exist (APA, 2006a; Elman, Illfelder-Kaye, & Robiner, 2005; Johnson & Campbell, 2002; 2004; Lichtenberg et al., 2007; Roberts et al., 2005). This has resulted in evaluation criteria that favor students who are smart and who are likely to fulfill the academic requirements of graduate training, but who may or may not possess the prerequisite interpersonal qualities that are generally accepted as necessary for success in the profession.

In sum, the central problem with early evaluation in psychology graduate programs is that it is overly focused on selecting and evaluating “smarts,” and does not adequately attend to ensuring that individuals become talented (or competent) practitioners. Many

psychologists over the course of history have hypothesized about which skills trainees need to have at the outset of training. This amalgamation of characteristics generally includes personality traits (e.g., openness, flexibility, tolerance), interpersonal qualities or skills (e.g., empathy, warmth), and emotional stability/health (Hatcher & Lassiter, 2007). Defining and adequately measuring the complex and ambiguous (but essential) skills, traits, and characteristics needed for graduate study in professional psychology has been an enormous challenge. However, there may be a way to solve this fundamental problem:

If we stop thinking of professional psychology as an arcane art and examine it instead as a set of complex professional occupations, the well-tried methods for measuring and improving performance that industrial/organizational psychologists have developed over the years can be applied to practitioners in our own field (Peterson, 2003, p. 797).

Along these lines, such methods could be instrumental in identifying individuals who are well-suited to the task of professional training in psychology, by virtue of not only their cognitive skills but also their interpersonal skills. Appropriate assessment methodologies could also be applied to evaluate these skills throughout practical training. The extensive literature on psychotherapy provides a strong empirical base from which to select variables that bear directly on the tasks of a professional psychologist and is central to informing what essential and relevant skills promising trainees should possess in addition to intellectual ability.

Empirical Directions for Identifying Promising Professionals

Research involving the practice of psychotherapy has resulted in a massive body of evidence supporting the assertion that psychotherapy works (Lambert & Ogles, 2004; Nathan & Gorman, 2002), as well as a great deal of evidence to suggest what makes it work or not (Castonguay et al., 2010; Hill, 2001; Lambert & Ogles, 2004; Orlinsky et al., 2004), and who makes it work or not (Clarkin & Levy, 2004; Beutler et al., 2004).

Although much remains to be discovered and fully understood about the complexities of how “psychotherapy works” (Castonguay et al., 2010; Lambert & Ogles, 2004; Orlinsky et al., 2004), this empirical body of evidence provides clear directions for identifying the skills that are most relevant for trainees to possess at the outset of training.

Psychotherapy: How does it Work?

The term “psychotherapy” is broad and can be used to refer to a multitude of activities conducted by a number of different individuals. Specifically, there are multiple “types,” “modes,” and “practitioners” of psychotherapy. The term psychotherapy will be used here to refer to an activity involving two or more individuals with the purpose of resolving some psychological problem or concern. At least one individual participant involved in the activity of psychotherapy, the therapist (or clinician, practitioner), is concerned with providing help for at least one other participant, the client (or patient), who is the individual seeking help. The term psychotherapy will be used to refer broadly to the multiple types of therapy and to the multiple modes in which psychotherapy occurs (e.g., individual, group, family); however the research reviewed here is largely focused on professional therapists providing individual therapy.

Decades of researchers have attempted to delineate what makes psychotherapy effective, but have not yet been able to agree on the mechanism by which psychotherapy works (for a review see Lambert & Ogles, 2004), a debate which has been referred to as the “great psychotherapy debate” (Wampold, 2001). Research on the mechanism of psychotherapy provides clues about what *causes* clients to improve. In turn, these factors help define the skills that individual practitioners need to be effective therapists. In other words, therapists (trainees) who provide more of what leads to client improvement will be more effective therapists (trainees) than those who provide less. The factors that lead to client improvement have been divided into “specific ingredients” that are unique to each type of therapy and “common ingredients” that are shared by most or all forms of psychotherapy.

Specific Ingredients. There are a multitude of different types of therapy that stem from divergent theories about how psychological problems develop and how they should be “treated” (e.g., cognitive-behavioral, psychodynamic, humanistic). Each type of therapy provides a specific explanation for what will help the client improve his or her functioning and how or why it will help. Therapists working within a particular model use specific techniques or “interventions” prescribed by the theory to facilitate improvement. These techniques are collectively referred to as “specific ingredients” because they are designated by a particular theory and are thought to influence clients (with a specific “mental disorder”) through a unique mechanism of change, thereby reducing symptoms of psychological suffering. According to the specific ingredients view, these model-specific techniques, or treatment packages result in positive client progress (Wampold, 2001).

The specific ingredients view has historically dominated the psychotherapy literature and has dictated the methodological details of studies designed to uncover which treatments are the most “efficacious” (Wampold, 2001) and has informed efforts to prescribe particular therapies for particular mental illnesses (Chambless et al., 1998; Task Force on Promotion and Dissemination of Psychological Procedures, 1995). A massive number of randomized clinical trials have measured the specific effects of particular therapies (standardized in therapy manuals) for certain disorders when compared to alternative or no treatment conditions. The randomized clinical trial design has also been used to isolate the effectiveness of specific techniques prescribed by specific therapies (Lambert & Ogles, 2004; Nathan & Gorman, 2002; Wampold, 2001).

Although research has overwhelmingly focused on the specific ingredients of psychotherapy, several outcome studies (e.g., Elkin et al., 1989; Imber et al., 1990; Project Match Research Group, 1998; Shapiro, Barkham, Rees, Hardy, Reynolds, & Startup, 1994; Sloane, Staples, Cristol, Yorkston, & Whipple, 1975) and meta-analytic reviews (e.g., Ahn & Wampold, 2001; Benish, Imel, & Wampold, 2008; Beutler, 1979; Cuijpers, van Straten, Andersson, & van Oppen, 2008; Grissom, 1996; Luborsky, Rosenthal, Diguer, Andrusyna, Berman, Levitt et al., 2002; Wampold, Mondin, Moody, Stich, Benson, & Ahn, 1997) suggest that when specific therapies are compared to one another, equivalent outcomes occur across clients and client problems. Small statistical advantages of cognitive-behavioral approaches have been attributed to experimenter allegiance effects (e.g., Gaffan, Tsauousis, & Kemp-Wheeler, 1995; Luborsky, Diguer, Seligman, Rosenthal, Krause, Johnson et al., 1999; Shapiro & Shapiro, 1982).

In contrast, there is evidence to suggest that certain cognitive-behavioral and behavioral approaches have advantages with certain problems, particularly severe psychological problems, such as obsessive-compulsive disorder, and populations, such as children, which suggests that specific ingredients under certain circumstances do have differential effectiveness (Lambert, Garfield, & Bergin, 2004). Along these lines, it has been argued that it is “hasty to conclude that there are no meaningful differences” among therapies (Chambless & Ollendick, 2001, p. 704; see also Chambless, 2002; DeRubeis, Brotman, & Gibbons, 2005; Kazdin, 2005, 2007). For instance, Chambless and Ollendick (2001) noted that behavioral therapies are the “treatments of choice” (p. 704) for severe problems.

Although this debate has not yet been settled, there is not much compelling support for the explanation that psychotherapy works *solely* because of the specialized procedures prescribed by a particular theory (Lambert & Archer, 2006; Lambert & Ogles, 2004), and hence specialized and technical therapist skills. Stated another way, specific ingredients do not seem to account for much of the outcome variance in psychotherapy, given that clients with similar problems improve as a result of different techniques (Lambert & Archer, 2006; Lambert, Garfield, & Bergin, 2004; Lambert & Ogles, 2004; Wampold, 2001; Wampold & Bhati, 2004). Equivalence among specific therapies implies that something other than specific ingredients, or something specific yet to be discovered, is accounting for more variance in psychotherapy outcomes. The explanation that has received the most attention, the common factors hypothesis, is the notion that all therapies work through similar mechanisms that account for more outcome variance than the specific techniques being employed (Lambert & Ogles, 2004).

Common Ingredients. According to the common factors hypothesis, there are elements of therapy, or common ingredients present in all psychotherapy activities. These common elements have often been treated as nonspecific, or *inactive* in the specific ingredients approach to psychotherapy; however, proponents of the common factors hypothesis assert that these ingredients are not “inert” or “trivial” (Wampold, 2001). Instead, common ingredients are central to the mechanism of psychotherapy, but do not revolve around specific theories (Butler & Strupp, 1986; Critelli & Neumann, 1984; Frank & Frank, 1991; Garfield, 1973; Lambert & Ogles, 2004; Luborsky et al., 2002; Luborsky, Singer, & Luborsky, 1975; Parloff, 1986; Stiles, Shapiro, & Elliott, 1986; Wampold, 2001).

Rosenzweig (1936/2002) provided the first description of how the common ingredients operate in psychotherapy, with his seminal claim: “At last the Dodo said, ‘*Everybody* has won, and *all* must have prizes’” (p. 5). This “dodo bird verdict” was based on the early observation that very different, conflicting therapy procedures led to effective change. Rosenzweig (1936/2002) asserted that there was a logical problem with the notion that contradictory specialized techniques caused the same client changes:

If such theoretically conflicting procedures...can lead to success, often even in similar cases, then therapeutic result is not a reliable guide to the validity of theory...It takes but little reflection to arrive at the roots of the difficulty from the standpoint of logical deduction...the same conclusion cannot follow from opposite premises...(p. 5).

Rosenzweig (1936/2002) suggested that the solution to this logical problem could be discovered in the elements common among all effective therapies: “when such a

contradiction appears...it is justifiable to wonder...whether the factors that actually are operating in several different therapies may not have much more in common than have the factors alleged to be operating” (p. 6).

Rosenzweig (1936/2002) asserted that psychotherapy outcomes were more likely the result of commonalities among therapies that had not been recognized. He proposed that these common ingredients were the “active” or “curative” ingredients in psychotherapy. Since this seminal paper, a wide array of common ingredients have been proposed (Beutler, 1983; Brady, Davison, Dewald, Egan, Fadiman, Frank et al., 1980; Garfield, 1973; Goldfried, 1980; Grencavage & Norcross, 1990; Karasu, 1986; Strupp, 1973) and there are several models of psychotherapy that are based on the common factors hypothesis (Frank & Frank, 1991; Garfield, 1995; Henry & Strupp, 1994; Hill, 2005; Lambert & Ogles, 2004; Orlinsky & Howard, 1986; Strupp, 1973; Wampold, 2001).

Taken together, common ingredient explanations typically include important characteristics, attitudes and expectations of the therapy *participants* (therapist and client) and the *relationship* that develops between them, as opposed to the specific ingredients’ sole emphasis on techniques (Frank & Frank, 1991; Hill, 2005; Lambert & Barley, 2002; Lambert & Ogles, 2004; Orlinsky & Howard, 1986; Strupp, 1973, 1977; Wampold, 2001). Consistent with the notion that common ingredients are “active,” the relationship that develops between the therapist and client has received a great deal of empirical attention and support (Norcross, 2002). The therapy relationship is universally included in common factor models and is at the very center of several of them (Frank & Frank, 1991; Henry & Strupp, 1994; Hill, 2005; Orlinsky et al., 2004; Wampold, 2001).

Though controversial (see Beutler, 2002; Chambless, 2002; Hrobjartsson & Gotzsche, 2007; Hunsley & Westmacott, 2007), several lines of converging evidence suggest that common ingredients are important to determining psychotherapy outcomes. First, contradictory therapy techniques result in positive client outcome. Second, the “placebo” effect suggests that when clients are randomly assigned to a “placebo” treatment control group (e.g., “supportive” therapy) they improve despite the absence of specific techniques (Baskin, Tierney, Minami, & Wampold, 2003; Grissom, 1996; Lambert, 2005; Lipsey & Wilson, 1993; Wampold, 2001; Wampold, Imel, & Minami, 2007; Wampold, Minami, Tierney, Baskin, & Bhati, 2005). Finally, a great deal of research has demonstrated that the relationship between the client and therapist has an important influence on psychotherapy outcomes (for a review see Norcross, 2002).

It seems clear that common factors exert a “substantial” influence on client outcomes, and their effects on the process of psychotherapy have been widely recognized, whereas the link between specific ingredients and client outcomes is arguably tenuous (Lambert & Ogles, 2004, p. 172; see also Lambert & Barley, 2002; Lambert, Garfield, & Bergin, 2004; Wampold, 2001, 2007). Consistent with this, several leading theorists and researchers characterize psychotherapy as an “interpersonal context” in which the effectiveness of specific techniques depends on the strength of the relationship between therapist and client (APA, 2006b; Barber, 2009; Beutler, 2002; Beutler & Harwood, 2002; Butler & Strupp, 1986; Goodheart, Kazdin, & Sternberg, 2006; Hatcher & Barends, 2006; Henry & Strupp, 1994; Hill, 2005; Kazdin, 2005, 2007, 2008, 2009; Lambert & Ogles, 2004; Messer, 2004; Norcross, 2002; Orlinsky et al., 2004; Wampold,

2001, 2007). Along these lines, research has begun to focus on the contribution of therapy participants to the therapy relationship as a means for improving outcomes.

The Role of the Therapy Relationship and Therapist Contributions

Empirical research on the role of the therapy relationship and individual (therapist and client) contributions to the relationship indeed indicate that relational factors play an important role in determining the success of psychotherapy (APA, 2006b; Lambert & Ogles, 2004). In fact, the therapy relationship is the strongest predictor of psychotherapy outcomes identified to date (Norcross, 2002; Orlinsky et al., 2004). Research suggests that individual therapists' (trainees') interpersonal contributions to the relationship significantly influence whether psychotherapy will be effective or ineffective (Lambert & Baldwin, 2009). Highly relevant to delineating which trainee skills might facilitate positive changes in clients and determining what skills are important to evaluate throughout training is empirical research that informs what therapists should do (or shouldn't do) to build a high quality therapy relationship.

What is the Therapy Relationship? Although psychologists generally agree that the therapeutic relationship is a common factor affecting change across different types of therapies, there is no gold standard definition of the relationship (Elvins & Green, 2008; Gaston, 1990; Horvath, 2006; Horvath & Bedi, 2002; Horvath & Luborsky, 1993). Contemporary research favors a pantheoretical view of the therapy relationship, informed by Bordin's seminal (1979) tripartite model of the "working alliance." Bordin (1979) posited that all therapies contained the working alliance, which includes three components. First, agreement about the "goals" of therapy involves the extent to which participants agree about what will lead to the desired outcome. Second, collaboration

about the “tasks” of therapy reflects the extent to which the therapist and client agree about what each participant needs to do to lead to the desired change. Finally, the “bond,” or the “nature of the human relationship” between participants, reflects feelings of affiliation (rapport), trust, and emotional attachment.

Most current constructions of the alliance (commonly referred to as the therapeutic, helping, or working alliance) contain Bordin’s (1979) basic elements. Several different alliance scales are commonly used in empirical research (Elvins & Green, 2008; Hatcher & Barends, 2006; Horvath, 2006; Horvath & Bedi, 2002) and, taken together, measure four dimensions: 1) client and therapist bond, or emotional connection (liking, respect, caring), 2) client positive engagement in therapy (or lack thereof), 3) therapist positive (or negative) contributions and involvement in therapy, and 4) “confident collaboration” or agreement on and belief in therapy and one another (Gaston, 1990). The term “alliance” will be used in this document to refer to “the quality and strength of the collaborative relationship between client and therapist” (Horvath & Bedi, 2002, p. 41). A great deal of extant research has focused on the association between the alliance and psychotherapy outcomes.

Empirical Association between Relationship Quality and Outcome. Recently, the chair of the task force on “empirically supported therapy relationships” concluded, “It’s the relationship stupid!” (Norcross, 2002, p., 5). This conclusion is supported by a massive literature base suggesting that strong alliances are related to better psychotherapy outcomes, whereas weak alliances are related to early drop out and poorer outcomes (Beutler et al., 2004; Castonguay & Beutler, 2006; Castonguay, Constantino, & Grosse Holtforth, 2006; Constantino, Castonguay, & Schut, 2002; Crits-Cristoph, Connolly

Gibbons, & Hearon, 2006; Hatcher & Barends, 2006; Horvath, 2006; Lambert & Barley, 2002; Lambert & Ogles, 2004; Luborsky, 1994; Orlinsky et al., 2004; Safran & Muran, 2006; Samstag, 2006). For instance, in a review of 109 studies (that yielded more than 1,000 separate findings), it was asserted that relational process variables (e.g., collaboration, bond, empathy, warmth) positively influence therapy outcomes with “few findings in this or related fields [that] seem better documented” (Orlinsky et al., 2004, p. 345).

Consistent with narrative reviews, the results of three meta-analytic reviews including up to 90 studies of individual therapy provided by experienced therapists in adult clinical settings (client $n > 5$) suggest that there is a modest association (r_w range = .21 - .26) between the strength of the alliance and client outcome across a wide variety of therapies and client concerns (Horvath & Bedi, 2002; Horvath & Symonds, 1991; Martin, Garske, & Davis, 2000). Additionally, meta-analytic reviews are consistent with more recent individual studies that continue to support the positive association between alliance and outcome (Botella, Corbella, Belles, Pacheco, Herrero, Ribas et al., 2008; Dinger, Strack, Leichsenring, Wilmers, & Schauenburg, 2008; Hoffart, Borge, Sexton, & Clark, 2009; Hoffart, Sexton, Nordahl, & Stiles, 2005; Klein, Schwartz, Santiago, Vivian, Vocisano, Castonguay et al., 2003; Missirlian, Toukmanian, Warwar, & Greenberg, 2005; Strauss, Hayes, Johnson, Newman, Brown, Barber et al., 2006; Zuroff & Blatt, 2006).

Despite the robust empirical association between alliance and outcome, researchers are divided on what this association *means*. Specifically, several researchers have concluded that the therapy relationship is important and have suggested that the relationship plays a universal *causal* role (Norcross, 2002), whereas others have remained

skeptical of its importance altogether (DeRubeis & Feeley, 1990; Feeley, DeRubeis, & Gefland, 1999). For instance, questions remain about whether the association between alliance and outcome is causal or instead represents a spurious statistical result that is related to early symptom changes (Barber, 2009; Crits-Christoph et al., 2006; DeRubeis et al., 2005; Kazdin, 2007). However, there is evidence that directly contradicts this assertion. For instance, several studies suggest that the strength of the early alliance (before symptom changes occur) is consistently related to “distal outcomes” (Beutler et al., 2004, p. 288; see also Barber, Connolly, Crits-Christoph, Gladis, & Siqueland, 2000; Klein et al., 2003; Suh, Strupp, & O’Malley, 1986).

Additionally, it is notable that the alliance does not account for but a small portion of the total outcome variance in psychotherapy (Beutler & Harwood, 2002; Crits-Christoph et al., 2006). Several researchers have suggested that the quality of the relationship may provide a filter through which specific interventions and techniques are received and perceived as helpful or unhelpful. That is, both techniques and alliance are important and each is less meaningful (i.e., decreased predictive power) when viewed in isolation. Along these lines, Barber (2009) noted: “undoubtedly, creating a good working relationship is an important therapeutic task...possibly a prerequisite” and may serve as a “thermometer” for how well the work is going (p. 3; see also Barber, Gallop, Crits-Christoph, Barrett, Klostermann, McCarthy et al., 2008; Castonguay, Goldfried, Wiser, Raue, & Hayes, 1996; Eaton, Abeles, & Gutreund, 1993; Fitzpatrick, Stalikas, & Iwakabe, 2001; Gaston, Piper, Debbane, Bienvenu, & Garant, 1994; Gaston, Thompson, Gallagher, Cournoyer, & Gagnon, 1998; Goldman, Greenberg, & Angus, 2006; Hill,

2005; Missirlian et al., 2005; Orlinsky et al., 2004; Rector, Zuroff, & Seagal, 1999; Strupp, 1958, 1998; Trepka, Rees, Shapiro, Hardy, & Barkham, 2004).

Thus, although questions remain regarding the nature of the link between alliance and outcome (Barber, 2009; DeRubeis, Brotman, & Gibbons, 2005; Kazdin, 2005, 2008) “reviewers are virtually unanimous in their opinion that the therapist-patient relationship is critical for positive outcome” (Lambert & Okiishi, 1997, p. 67; see also Beutler et al., 2004; Binder & Strupp, 1997; Castonguay & Beutler, 2006; Henry & Strupp, 1994; Norcross, 2002; Orlinsky et al., 2004). The alliance, as a robust predictor of therapy outcomes involves both client and therapist contributions. Whereas clients significantly influence the alliance (Clarkin & Levy, 2004), training programs have little power over the types of clients that its graduates will interact with as professionals. However, graduate programs do have control over the therapists that are allowed entry into the profession, and evidence suggests that certain therapist (trainee) interpersonal skills have a substantial influence on therapy relationships and outcomes.

Therapist Relational Contributions. Psychotherapy is an interpersonal interaction between therapist and client, and “the inescapable fact of the matter is that the therapist is a person, however much he may strive to make himself an instrument of his patient’s treatment (Orlinsky & Howard, 1977, p. 567). This assertion is supported by decades of empirical research suggesting that the therapist significantly influences the therapy relationship and outcome (Aveline, 2005; Bergin, 1997; Beutler, 1997; Beutler et al., 2004; Gurman & Razin, 1977; Henry & Strupp, 1994; Kiesler, 1996; Krause & Lutz, 2009; Krause, Lutz, & Saunders, 2007; Lambert, 1989, 2007; Lambert & Baldwin, 2009; Lambert & Okiishi, 1997; Orlinsky et al., 2004; Rogers, 1957; Strupp, 1958, 1995, 1998;

Teyber & McClure, 2000; Wampold, 2001). This notion similarly applies to trainees, who bring their interpersonal skills to the training table, skills which influence trainee effectiveness.

Indeed, a great deal of evidence suggests that some therapists (trainees) consistently offer *more* of what leads to client improvement, whereas some therapists (trainees) offer *less* of what leads to client improvement across a wide range of individual clients and problems with varying levels of severity (e.g., Blatt, Sanislow, Zuroff, & Pilkonis, 1996; Huppert, Bufka, Barlow, Gorman, & Shear, 2001; Kim, Wampold, & Bolt, 2006; Luborsky, Crits-Christoph, McLellan, Woody, Piper, Liberman et al., 1986; Luborsky, McLellan, Woody, O'Brien, & Auerbach, 1985; Lutz, Leon, Martinovich, Lyons, & Stiles, 2007; McLellan, Woody, Luborsky, & Goehl, 1988; Okiishi, Lambert, Eggett, Nielsen, & Dayton, 2006; Okiishi, Lambert, Nielsen, & Ogles, 2003; Orlinsky & Howard, 1980; Project Match Research Group, 1998; Shapiro, Firth-Cozens, & Stiles, 1989; Yalom & Lieberman, 1971; for reviews see Crits-Christoph, Branacki, Kurcias, Beck, Carroll, Perry et al., 1991; Crits-Christoph & Mintz, 1991; Lambert, 1989; Luborsky, McLellan, Diguier, Woody, & Seligman, 1997). Notably, preliminary evidence suggests that differences among therapists are stable over time (Brown, Lambert, Jones, & Minami, 2005; McLellan et al., 1988; Wampold & Brown, 2005).

Individual variation in therapist effectiveness does not seem related to technical skill, experience, training level, or other demographic characteristics (Beutler et al., 2004; Huppert et al., 2001; Lafferty, Beutler, & Crago, 1989; Lambert, 1989; Najavits & Strupp, 1994; Okiishi et al., 2006; Okiishi et al., 2003; Wampold & Brown, 2005). Instead, therapists with particularly effective relational stances (e.g., warm, supportive,

accepting, empathic) reliably perform better than those who are interpersonally *ineffective* (e.g., hostile, critical, neglectful, blaming, and controlling; Anderson, Ogles, Patterson, Lambert, & Vermeersch, 2009; Baldwin, Wampold, & Imel, 2007; Binder & Strupp, 1997; Castonguay et al., 2010; Dinger et al., 2008; Henry, Schacht, & Strupp, 1986, 1990; Henry, Schacht, Strupp, Butler, & Binder, 1993; Henry & Strupp, 1994; Henry, Strupp, Butler, Schacht, & Binder, 1993; Jennings & Skovholt, 1999; Lafferty, Beutler, & Crago, 1989; Luborsky et al., 1997; Najavits & Strupp, 1994; Suh, Strupp, & O'Malley, 1986).

Thus, extant literature suggests that “the therapist’s ability to form an alliance is possibly the most crucial determinant of his effectiveness” (Luborsky et al., 1985, p. 610). This assertion is consistent with theoretical views of the therapist as 1) an attachment-like figure, who effectively helps clients change by providing a caring, protective, secure base (i.e., a good parent; Bowlby, 1988; Henry & Strupp, 1994; Lampropoulos, 2001; Obegi, 2008; Strupp, 1973, 1977), 2) a persuader who exerts positive influence through confident expertise and skillful verbal communication (Frank & Frank, 1991; Johnson & Matross, 1977; Strupp, 1973, 1977; Strong, 1968), and 3) a teacher-like figure who models adaptive thinking, behaviors, and social skills (Henry & Strupp, 1994; Strupp, 1973, 1977; Wilson & Evans, 1977).

If an important part of the professional therapist’s job is to be a relational figure, it follows that trainees should come to graduate school at least partially equipped to effectively fulfill this relational role, in addition to being intellectually equipped to learn specialized techniques and skills. Notably the relational skills that are needed to be an effective professional are not mystical therapist powers attained during graduate school,

but are the same interpersonal skills that lead to high quality, healthy social relationships. Interpersonal skills (like intellectual abilities) do not begin developing during graduate training. Rather, these skills develop over the course of an individuals' lifetime and can be refined or enhanced throughout training (Berk, 2000; Bowlby, 1988; Feshbach, 1997; Rogers, 1957; Vivino, Thompson, Hill & Ladany, 2009; Watson, 2002). Thus, it seems ideal to delineate what baseline interpersonal skills a promising trainee might possess, and attempt to measure the presence or absence of these skills in professional trainees.

Indeed, it seems likely that trainees who possess these important relational skills at the outset of professional training will be more likely to master them as they are learning to become therapists. Although very little empirical research has attempted to identify promising trainees on the basis of these interpersonal skills, or to capture these skills reliably and meaningfully, a whole host of skills, traits, and propensities are thought to contribute to the quality of the therapy relationship and client outcomes. Research provides clear directions for determining which core interpersonal skills relate to client outcomes. Trainees who possess the basic interpersonal skills described below are likely to be able to build helpful therapy relationships and facilitate better outcomes with clients.

Therapist Interpersonal Skills and Clinical Effectiveness

Although a large number of therapist characteristics (e.g., personality traits, demographics, and attitudes) could conceivably contribute to the quality of the therapy relationship (Beutler et al., 2004), interpersonal skills demonstrated through a series of moments in the context of therapy-like interactions (simulated or real) have the most

robust association with client outcomes (Anderson et al., 2009; Asay & Lambert, 2002; Beutler et al., 1994; Binder & Strupp, 1997; Lambert & Baldwin, 2009; Lambert & Okiishi, 1997; Norcross, 2002). Stable traits and characteristics undoubtedly influence how therapists behave interpersonally; however the remainder of this document is focused on conceptualizing interpersonal skills that are “grounded and expressed in the emerging realities of the therapy session” (Horvath & Bedi, 2002, p. 56). In other words, a trainee’s interpersonal ability with a spouse may be related to the trainee’s overall relational ability as a human, but performance in the context of therapy is the most relevant to judging therapist effectiveness.

Essentially, the interpersonal skills that seem most clearly and consistently related to better clinical performance in the empirical literature can be boiled down to “the capacity to express sensitivity to the client’s needs,” or empathic responsiveness, “the ability to generate a sense of hope,” or a caring, warm, and hopeful stance, and the “ability to respond to challenges,” or manage interpersonal conflict in the therapy relationship (Horvath & Bedi, 2002, p. 56-57; see also Norcross, 2002). These therapist skills were included in a recent review of significant therapist contributions to the therapy relationship, organized by the task force on empirically supported therapy relationships (Norcross, 2002). The task force concluded that many therapist interpersonal skills influence client outcomes and should therefore be explicitly considered in training and practice.

The three skills (empathy, warmth, and ability to negotiate interpersonal conflict) outlined in this document are basic interpersonal skills that seem likely to be components of the capacity to be a successful (i.e., clinically competent) professional. Although these

skills are statistically and conceptually interrelated (Barrett-Lennard, 1962; Blatt, Zuroff, Quinlan, & Pilkonis, 1996; Bohart, Elliott, Greenberg, & Watson, 2002; Gurman, 1977; Horvath & Greenberg, 1986; Rogers, 1957; Truax & Carkhuff, 1967), each will be discussed separately in service of clarity.

Empathy

Empathy is perhaps the most empirically supported therapist interpersonal skill in the psychotherapy literature (Rogers, 1975). Therapists who are judged as more empathic by their clients and outside observers often build stronger, more therapeutic relationships with clients and globally have better client outcomes (Ackerman & Hilsenroth, 2003; Asay & Lambert, 2002; Bohart et al., 2002; Gurman, 1977; Patterson, 1984; Strupp, 1998; Teyber & McClure, 2000; Truax & Carkhuff, 1967; Watson, 2002).

What is empathy? Empathy has been described as “the accepting, confirming, and understanding human echo ... a psychological nutrient without which human life as we know and cherish it, could not be sustained” (Kohout, 1978, p. 705 as cited in Barrett-Lennard, 1993). As such, empathy has been referred to as the “basis of all human interaction” (Duan & Hill, 1996, p. 262). Empathy has been conceptualized in several different ways (Bohart et al., 2002; Bohart & Greenberg, 1997b; Eisenberg, Fabes et al., 1991; Feshbach, 1997). Historically, empathy has been abstractly defined as “reading or feeling” into another person (Titchener, 1924 as cited in Duan & Hill, 1996, p. 262), sensing another persons thoughts and feelings “as if they were your own” (Rogers, 1957/2007, p. 243), or standing in someone else’s shoes (Katz, 1963). Empathy has been conceptually defined as a stable disposition or ability, a context or situation-specific state,

and an unfolding process (Duan & Hill, 1996; see also Barrett-Lennard, 1962, 1981, 1993).

According to the most recent conceptualizations, empathy is a multifaceted construct that includes four dimensions (Bohart et al., 2002; Bohart & Greenberg, 1997a, 1997b; Watson, 2002). First, the cognitive dimension is an ability to accurately understand the perspective of another, or “intellectual empathy” (Duan & Hill, 1996). Second, the affective dimension involves being emotionally “attuned,” and reflects awareness of and concern about the client’s needs, or “empathic emotions” (Duan & Hill, 1996). This dimension includes the therapist’s ability to experience and express compassion, the desire to help, warmth, and hopefulness for clients (Vivino et al., 2009). Third, empathy involves a response (verbal or nonverbal) that clearly communicates expert understanding, attunement, and a focus on the client’s immediate expressions, thoughts, and feelings. Finally, empathy involves “a way of being together in relationship” (Bohart & Greenberg, 1997, p. 419), or “stepping into” the client’s world and relating to the client, using their language (Bohart & Greenberg, 1997; Rogers, 1957).

Consistent with a multi-dimensional view of empathy, the term will be used here to refer to an unfolding interpersonal process (Barrett-Lennard, 1981, 1993; Bohart et al., 2002; Bohart & Greenberg, 1997; Vanaerschot & Lietaer, 2007). Although ratings of therapist empathy are associated with particular verbal (e.g., advice, interruptions; Barkham & Shapiro, 1986) and nonverbal (e.g., eye contact, facial expression; Tepper & Haase, 1978) behavior, current notions suggest that empathy cannot be meaningfully captured by a simple verbal response (Barrett-Lennard, 1981; Bohart et al., 2002; Bohart & Greenberg, 1997b; Lambert, DeJulio, & Stein, 1978; Watson, 2002). That is, empathy

involves “stepping into” each unique client’s world, and can be thought of as an ongoing effort to learn and speak another person’s language, rather than a predetermined, “cookie cutter” response. Along these lines, research suggests that what is empathic for one client is not necessarily empathic for another client (Bachelor, 1988).

Thus, empathy is an emotionally responsive (listening, attending, sensitive and caring) manner of interacting in relation to the client as they express (implicitly or explicitly) their needs, as well as the ability to articulate and express this sensitive, caring, connected expert understanding (Bohart et al., 2002; Rogers, 1957) in a manner that is focused on the client. Empathy may take many forms depending on the needs of the client such that “all good therapy responses should be conveyed empathically” (Bohart & Greenberg, 1997b, p. 431). For instance, an empathic therapist might remain silent after sensing that a client needs time or interpersonal distance. Alternatively, a therapist might empathically “reflect” what the client is feeling or make an empathic interpretation (Bohart et al., 2002; Hill, 2004). In contrast, unempathic therapist responses are expressed in a neglectful, dismissive, disconnected, inattentive, or inflexible manner.

There is reason to believe that therapist empathy is not a specialized professional ability, but an interpersonal skill that trainees already possess at the outset of training (Feshbach, 1997; Hatcher & Lassiter, 2007; Rogers, 1957). Indeed, empathy “is a basic form of social communication” that has been measured as early as infancy, and is referred to as theory of mind in children (Feshbach, 1997, p. 33; see also Berk, 2000; Eisenberg & Fabes, 1990). Although it seems likely that professional therapists learn to use empathy in a more intentional, purposeful, skilled, and frequent manner in their

interactions with clients throughout training, therapeutic empathy seems to be made up of the “same stuff” as empathic communication in all positive relationships, including the therapy relationship.

Empathy and Client Outcome. A large body of empirical literature supports the notion that more empathic therapists are more effective with their clients (Ackerman & Hilsenroth, 2001, 2003; Asay & Lambert, 2002; Bohart et al., 2002; Bohart & Greenberg, 1997; Gurman, 1977; Luborsky, Chandler, Auerbach, Cohen, & Bachrach, 1971; Norcross, 2002; Orlinsky et al., 2004; Patterson, 1984; Sexton, Littauer, Sexton, & Tommeras, 2005; Teyber & McClure, 2000; Truax & Carkhuff, 1967; Traux & Wargo, 1966). Indeed, empathy is a “demonstrably effective” element of the therapy relationship (Norcross, 2002). Although there are instances in the literature when empathy is weakly or unrelated to outcome (Lambert et al., 1978), the overwhelming majority of studies suggest that empathy is a key interpersonal skill, particularly when measured from the perspective of the client.

For instance, in a comprehensive review of psychotherapy process and outcome research, Orlinsky et al. (1994, 2004) found that 54% of the empirical associations ($n = 115$) between therapist empathy (“expressive attunement”) and client outcome were significantly positive and none were negative. When empathy was rated from the perspective of the client, 72% of the empirical associations ($n = 47$) were significantly positive (Orlinsky et al., 1994, 2004) suggesting a relatively robust link between empathy and client outcomes (see also Asay & Lambert, 2002; Elliott, Bohart, Watson, & Greenberg, 2011; Gurman, 1977; Patterson, 1984).

In a recent meta-analysis of 47 studies, empathy had a medium association with client outcomes ($r_w = .23 - .32$; Bohart et al., 2002; see also Elliott et al., 2011). This association remained modest across a wide variety of therapy approaches, including experiential, psychodynamic, and cognitive-behavioral therapies as well as client ($r = .25$) and observer ($r = .23$) perceptions of empathy (Bohart et al., 2002). Therapist ratings of their own empathy were consistently unrelated to outcome which may be due to the tendency for therapists to overestimate their abilities (Bohart et al., 2002; Dooley, 1975; Gurman, 1977; Kurtz & Grummon, 1972). Notably, empathy was a stronger predictor in therapy with *less* experienced therapists, implying that the importance of empathy decreases as therapists add more technical skills to their repertoire (Bohart et al., 2002).

Additionally, qualitative research involving therapy clients' perspectives about what is and what is not helpful in therapy strongly suggest that empathy is a universally important piece of what makes psychotherapy helpful (Bachelor, 1988, 1995; Bedi, 2006; Elliott & James, 1989; Thompson & Hill, 1993; Timulak, 2007). Notably, one qualitative study suggests that "one size does not fit all" in terms of what empathy is or how it is helpful (Bachelor, 1988, 1995). For instance, in one sample of clients ($n = 52$), some described cognitive empathy (44%), whereas others described affective empathy (30%; Bachelor, 1988). The client's perspective in this study clearly suggests that empathy is not mechanical or captured meaningfully by intellectual or verbal response modes, but dependent on a therapists' ongoing sense of what each client needs.

Finally, studies that have directly contrasted more effective with less effective therapists and their differential impact on client outcomes suggest that more effective therapists are significantly more empathic (Anderson et al., 2009; Barret-Lennard, 1962;

Lafferty, Beutler, & Crago, 1989; Luborsky et al., 1985; Najavits & Strupp, 1994; Strupp, 1998; Truax & Carkhuff, 1967). For instance, Najavits and Strupp (1994) differentiated the most effective from the least effective professional therapists (who were all “highly recommended” by professional colleagues; $n = 16$) on the basis of client ($n = 80$) outcomes (there were no differences among clients prior to treatment). Therapist in-session interpersonal behavior was rated on the basis of global observer process ratings. The most effective therapists in this study demonstrated more positive behaviors in session, including warmth, understanding (i.e., empathy), and protectiveness, and fewer negative behaviors including ignoring (i.e., lack of empathy), rejecting, or belittling (Najavits & Strupp, 1994). Along similar lines, in a sample of graduate student therapists ($n = 30$) and outpatient clients ($n = 60$), empathy was the strongest predictor of therapist effectiveness (Lafferty, Beutler, & Crago, 1989).

Taken together, the robust link between empathy and client outcome clearly suggests that empathy is a core interpersonal skill that psychotherapists should possess (Ackerman & Hilsenroth, 2003; Asay & Lambert, 2002; Bohart & Greenberg, 1997a; Bohart et al., 2002; Elliott et al., 2011; Norcross, 2002; Patterson, 1984; Rogers, 1957; Strupp, 1998; Teyber & McClure, 2000; Watson, 2002). As such, empathic responsiveness is emphasized in all therapy approaches (for reviews see Bohart et al., 2002; Watson, 2002). Thus, there is reason to believe that therapist empathy is an interpersonal skill that trainees should possess at a baseline level from the outset of training. However, very limited research has addressed whether empathic ability can be reliably measured or is predictive of clinical performance in trainees.

Warm Affiliation

In addition to empathic ability, “unconditional positive regard” (Rogers, 1957) has been viewed as “vital to the provision of empathy” (Watson, 2002, p. 446).

Unconditional positive regard has also been referred to as respect, warmth (or nonpossessive warmth), compassion, support, affirmation, approval, reassurance, acceptance, caring, prizing, interest, and liking (Farber & Lane, 2002; see also Orlinsky et al., 2004; Truax & Carkuff, 1967). In short, warmth (or warm affiliation) falls on one end of the interpersonal dimension of affiliation in general theories of interpersonal interaction (Benjamin, 1974; Henry, 1996; Keisler, 1996; Leary, 1957). From this perspective, warmth can be contrasted with hostility, blaming, and belittling, general interpersonal behaviors that have implications for any relationship.

Notably, warmth and empathy are highly interrelated. The two, however, are not inseparable. Specifically, a therapist could be expressing warmth and caring concern, but might simultaneously “miss” the client’s experience or feelings and be perceived as unempathic. For instance, a therapist might respond to the client’s breakup with a significant other by expressing caring concern (“You must be so devastated”), which would effectively communicate warmth, but at the same time could miss part of the client’s feeling or experience of the break-up (e.g., the client is actually relieved to be out of the relationship). It seems very unlikely that a genuinely empathic therapist could also be realistically perceived as uncaring (hence they are often interrelated). In contrast, it seems quite likely for a therapist to express warmth without being empathic (hence they can be separate) and research suggests that warmth, in and of itself, is important.

What is Warm Affiliation? Rogers (1957/2007) defined warmth as the “prizing” of another person, which included acceptance, a non-evaluative stance, and “a caring for the client...as a *separate* person” (p. 243). Rogers (1957) conceptualized warmth as a therapist stance or attitude toward the client that was consistently communicated in session, but noted that therapists would sometimes experience a “conditional positive regard” (evaluative) or a “negative regard” (hostility). In other words, even a therapist who is generally warm may at times have negative or evaluative feelings toward a client. On the other hand, Rogers (1957) asserted that hostility would be unlikely or infrequent in successful or helpful interactions.

Vivino and colleagues (2009) theory of “compassion” (based on qualitative interviews with 14 professionals nominated as compassionate) most closely represents what is meant by the terms “warmth” and “warm affiliation” used in this document. Warmth is defined here as “a state of being” (p. 167) that is “broader and deeper than empathy” (p. 167), and creates the intimate bond or connection between the therapist and the client. Warmth in this sense originates from the therapist’s concern for the client’s “suffering,” hopefulness that the client can feel better, and an authentic desire or motivation to help the client with their suffering (see also Farber & Lane, 2002).

Vivino and colleagues (2009) conceptualized warmth (or compassion) as a “precursor” to empathy. This suggests that warmth and caring concern for clients must come before empathy can be truly experienced and expressed. For example, empathy has often been defined as merely having an accurate understanding of another entity, even an inanimate object (e.g., a person flying a kite needs to know where the kite is headed to be a successful kite flyer) and has been implicated in crimes of sadistic violence, in which

the perpetrator understands what will hurt the victim most and uses this understanding to cause the utmost harm (Shlien, 1997). Clearly then, in the therapist role, warmth is part and parcel of genuine, therapeutic empathy.

Warmth does not necessarily imply that therapists cannot challenge or question clients. Indeed, therapist challenging in session can be the result of authentic, caring concern for the welfare of another. Along these lines, it has been noted that “it is possible to be challenging, even critical in a manner that is not interpersonally hostile” (Henry & Strupp, 1994, p. 69). However, responding in a cold, blaming, or careless manner has negative consequences for clients. Thus, therapist interpersonal warmth, much like empathy, is expressed as caring concern, a desire to help, hopefulness for and about the client’s ability to change, expressed throughout their interactions with clients, but does not necessitate against challenging interventions.

Like empathy, expression of warmth in relationships is a skill that trainees could be expected to possess prior to training (Hatcher & Lassiter, 2007). Compassion has recently been construed as an “innate capacity” that develops through relational experiences (Vivino et al., 2009, p. 162) and is conceptually related to theories of how adult attachment styles impact relationships (Berk, 2000; Bowlby, 1988; Daniel, 2006; Obegi, 2008). Relational experiences likely impact the therapists’ feeling and expression of warmth toward therapy clients, such that trainees come equipped with a relatively stable set of expectations which might or might not be consistent with expression of caring concern for clients (Henry, 1996). Indeed research suggests that therapists’ relational skills aren’t necessarily amenable to training (Henry, Schacht et al., 1993; Henry, Strupp et al., 1993).

Warmth and Client Outcome. A large literature suggests that therapist warmth is associated with better client outcomes, whereas therapist hostility is related to client deterioration or poor outcomes (Ackerman & Hilsenroth, 2001, 2003; Asay & Lambert, 2002; Barrett-Lennard, 1962, 1981; Binder & Strupp, 1997; Farber & Lane, 2002; Gurman, 1977; Mitchell, Bozarth, & Krauft, 1977; Norcross, 2002; Orlinsky et al., 2004; Patterson, 1984; Teyber & McClure, 2000; Truax & Carkhuff, 1967; Truax & Mitchell, 1971). Indeed, the task force on empirically supported therapy relationships concluded that therapist warmth was a “promising and probably effective” therapist contribution to the relationship (Norcross, 2002). This conclusion is supported by several decades of research.

In a recent review of the empirical association between therapist warmth and client outcome, effect sizes were often modest (Farber & Lane, 2002). These studies taken together, suggested that when all rater perspectives were considered, approximately half (49%) of the effect sizes ($n = 55$) indicated a significant positive association between warmth and client outcome. When the client’s perspective was examined exclusively, 83% of the findings suggested that therapist warmth significantly contributed to better client outcomes ($n = 12$; Farber & Lane, 2002). Unfortunately, many of the studies included in this review were based on indices of warmth gleaned from various alliance scales, such that the association between warmth and outcome is confounded with the broader construct of the alliance (Farber & Lane, 2002).

Nonetheless, direct examinations of the influence of therapist warmth (or hostility) on the therapy process provides compelling evidence for the importance of this skill in diverse therapy approaches (Anderson et al., 2009; Duff & Bedi, 2010; Klee, Abeles, &

Muller, 1990; Lafferty, Beutler, & Crago, 1989; Strupp & Hadley, 1979; Tasca & McMullen, 1992). For instance, in a frequently cited study that compared more effective to less effective therapists (Najavits & Strupp, 1994), ratings of warmth (based on an observer-rated global process scale) and “affirmation” (based on client ratings on a fine-grained interpersonal process measure) were significantly and positively related to the length of stay in psychodynamic therapy. Notably, therapist warmth was not significantly related to client outcomes based on indices of symptom change in this study. The results in this sample suggest that warmth may help clients remain in the therapy relationship (i.e., perhaps warmth is necessary), but also imply that warmth alone was not enough to facilitate client change (i.e., perhaps warmth is not sufficient).

On the other hand, therapist warmth in another sample was significantly associated with decreased client symptomatology. Specifically, in one of the largest randomized clinical trials conducted to date (NIMH Treatment of Depression Collaborative Research Program; Elkin et al., 1989), supervisor ratings of therapist responses were averaged over the first four sessions of interpersonal therapy (Rounsaville, Chevron, Prusoff, Elkin, Imber, Sotsky et al., 1987). Professional therapist ($n = 11$) expressions of warmth were significantly correlated with client ($n = 35$) outcomes ($r = .40 - .60$; Rounsaville et al., 1987). Overall, the findings in this study suggest that therapist warmth does facilitate client change, and has a relatively strong relationship with symptom changes.

Similarly, research suggests that higher levels of therapist hostility are associated with client deterioration or poor outcomes. For example, two seminal studies have examined the effect of therapist hostility on outcome by evaluating the intricate interpersonal interactions (referred to as “process”) of therapists and clients in poor vs. good outcome

cases (total dyads = 22; Henry, Schacht, & Strupp, 1986, 1990). Results based on two separate data-sets yielded the same conclusion (Binder & Strupp, 1979; Strupp & Binder, 1984). Specifically, these two studies revealed that poor outcome cases were clearly differentiated from good outcome cases by higher levels of negative interpersonal process (e.g., reciprocal hostility and domination) and complex messages (i.e., a positive message that simultaneously communicates hostility). In contrast, good outcome cases were characterized by a significantly higher frequency of therapist affirming, helping, and protecting and an absence of complex communications (Henry, Schacht, & Strupp, 1986, 1990). These studies not only suggest that therapist warmth impacts the therapeutic process, but also imply that hostility is detrimental and can be detected even when therapists attempt to hide their negative reactions.

The results of these two seminal studies have been replicated in several more recent studies based on a variety of different therapy approaches (e.g., Coady, 1991a, 1991b; Hilliard, Henry, & Strupp, 2000; Samstag, Muran, Wachtel, Slade, Safran, & Winston, 2008; von der Lippe, Monsen, Ronnestad, & Eilertsen, 2008; for reviews see Binder & Strupp, 1997; Constantino, 2000; Strupp, 1998). Taken together, this research provides relatively clear support for the conclusion that “even small amounts of ‘negative process’ can lead to poor therapeutic outcomes” (Strupp, 1998, p. 27; see also Binder & Strupp, 1997; Henry, 1996; Henry & Strupp, 1994), but unfortunately does not clarify whether warmth alone leads to symptom change (necessary vs. sufficient).

Finally, in a unique and serendipitous demonstration of the importance of therapist warmth, Strupp and Hadley (1979) set out to isolate the influence of specific therapy techniques vs. common relational therapy ingredients. In this study, college males

seeking therapy (n = 30) were randomly assigned to either experienced professional therapists (n = 5) or to untrained nonprofessional therapists (n = 5). Notably, nonprofessionals were male professors, specifically selected as “therapists” because they had a wide reputation for being warm, friendly, and interpersonally skilled. The bottom line finding in this seminal study was that nonprofessional therapists performed equally well when statistically compared to the seasoned, professional therapists (Strupp & Hadley, 1979). Warm nonprofessionals were able to effectively help clients without the advantage of “technical” intervention, meaning that nonprofessional outcomes in this study seem wholly attributable to therapists who influenced their clients “through interested and concerned listening...positive feedback, encouragement, support, and occasional direct advice” (Strupp, 1998, p. 25).

Thus, therapist warm affiliation and respect toward therapy clients is likely central to achieving positive therapy outcomes (Asay & Lambert, 2002; Binder & Strupp, 1997; Farber & Lane, 2002; Norcross, 2002). At minimum, warmth seems necessary for creating a climate in which clients feel comfortable enough to remain in therapy (e.g., Najavits & Strupp, 1994) but may also facilitate client symptom changes (Rounsaville et al., 1987). The absence of therapist warmth (e.g., hostility) has negative consequences for the therapeutic process (Binder & Strupp, 1997). Along these lines, “virtually all schools of therapy either explicitly or implicitly promote the value of this basic attitude” toward clients (Farber & Lane, 2002, p. 176). Overall, this research suggests that it is important for trainee therapists to have a basic ability to experience and demonstrate interpersonal warmth. However a paucity of research has attempted to assess interpersonal warmth in trainees and link this skill with trainee performance.

Therapist warm and empathic stances have received a great deal of attention in the psychotherapy literature and are associated with better client outcomes (Bohart et al., 2002; Farber & Lane, 2002; Norcross, 2002). Although these skills are crucial for psychotherapists to express consistently in interactions with clients, there is evidence to suggest that therapists who are warm and empathic will have difficulty maintaining these therapeutic stances during challenging interpersonal interactions (Binder & Strupp, 1997). The ability to respond effectively in the context of a challenging interpersonal interaction has received recent attention and seems promising for evaluating the extent to which therapists are skilled enough to maintain a therapeutic stance in the context of interpersonal conflict.

Negotiating Interpersonal Conflict

Interpersonal conflicts and misunderstandings inevitably arise between clients and therapists, just as they do in any intimate relationship. Indeed, according to recent research, “ruptures” in the therapeutic alliance, or negative interpersonal processes, represent a routine part of the unfolding interaction between the client and therapist in psychotherapy (Safran & Muran, 2000; Safran, Muran, Samstag, & Stevens, 2002). Research suggests that the therapists’ ability to recognize and manage negative processes partially determines the success of therapy (Ackerman & Hilsenroth, 2001; Binder & Strupp, 1997; Coutinho, Ribeiro, Hill, & Safran, 2011; Safran et al., 2002). Indeed, the most effective psychotherapists may be individuals who are “especially careful about maintaining their composure” and “make as their highest priority adherence to an empathic, respectful stance toward their patients” (Binder & Strupp, 1997, p. 124) even in the face of challenging interactions.

What is an Alliance Rupture? Ruptures in the alliance are reflective of disagreements in negotiating the tasks and goals of therapy, or strains in the bond between the client and the therapist. Safran and Muran (2006), have defined alliance ruptures as “breakdowns in the collaborative process,” “poor quality relatedness,” “deterioration in the communicative situation,” or “failure to develop a collaborative process from the outset” (p. 288; Safran, Crocker, McMain, & Murray, 1990; Safran & Muran, 2000; Safran Muran, & Samstag, 1994; Safran et al., 2002). For example, in a strain in the therapeutic bond, a client who feels misunderstood by their therapist may disengage from the work of therapy (or may have never engaged in the first place) without openly discussing his or her feelings of being misunderstood. Indeed the client may even feel angry or disappointed with the therapist.

Generally, clients respond differently to strains or problems in the therapeutic alliance. Thus, Safran and Muran (2000) have also distinguished between ruptures characterized by client withdrawal (e.g., distancing, implicit, covert) vs. client confrontation (e.g., hostile, explicit, overt). For instance, a client who feels misunderstood by the therapist might covertly withdraw by remaining unresponsive to therapist interventions or by agreeing with everything the therapist says. In contrast, another client who similarly feels misunderstood might respond by overtly expressing anger or hostility toward the therapist. This might take the form of direct confrontation about the problem, or the client may indirectly attack the therapists’ expertise, competence, or approach to the work (Safran & Muran, 2000).

Alliance ruptures involve both client and therapist contributions to the interpersonal process (Safran & Muran, 2006). That is, ruptures occur when therapist and client are

engaging in a negative interpersonal interaction that is reciprocal. There are several explanations based on different types of therapy regarding the “source” and “nature” of negative interpersonal process or ruptures (Binder & Strupp, 1997; Hill & Knox, 2009). For instance, Safran and Muran (2006) have asserted that ruptures are caused by client transference and therapist countertransference based on in session behaviors. Essentially, transference is reflective of client reactions to the therapist that stem from therapist in-session behavior as well as on client personal history. Similarly, therapist countertransference is representative of therapist reactions to the client, based on client in-session behavior as well as therapist personal history (Safran & Muran, 2006).

Thus, a rupture is a “breakdown” in the interpersonal interaction between client and therapist (Safran & Muran, 2000) or “an episode of covert or overt hostile sentiments, and often behavior that snares both patient and therapist” (Binder & Strupp, 1997, p. 123). Although ruptures involve both clients and therapists, only therapist contributions and reactions to alliance problems are relevant to the present review. Ruptures are a relatively routine part of the process in a wide variety of therapies and therapists should be able to successfully deal with conflict in the therapy relationship (Safran et al., 2002). Notably, ruptures are not always extreme (e.g., a client simply avoids an emotionally loaded topic), can occur regularly throughout therapeutic encounters, and can reflect subtle as well as dramatic shifts in the strength of the alliance. Several therapist interpersonal skills seem to contribute to effective negotiation of conflict in the therapy relationship.

Successful Resolution of Alliance Ruptures. Some have theorized that the extent to which a therapist is able to disentangle themselves from a problem in the relationship and

take steps to repair it determines whether the alliance will be able to “withstand” the problem, ultimately influencing the fate of therapy (Binder & Strupp, 1997; Henry & Strupp, 1994; Hill & Knox, 2009; Safran & Muran, 2000; Safran et al., 2002). From this perspective, successful negotiation of conflict in the relationship involves a cluster of interpersonal skills, including the ability to 1) recognize conflict, 2) tolerate internal and interpersonal conflict, 3) respond openly (non-defensively) in a warm, empathic manner, and 4) metacommunicate (Hill, Kellems, Kolchakian, Wonnell, Davis, & Nakayama, 2003; Hill & Knox, 2009; Safran & Muran, 2000; Safran, Muran, & Samstag, 1994; Safran et al., 2002).

Recognition of interpersonal conflict or ruptures in the alliance requires the therapist to “observe the process” (Binder & Strupp, 1997, p. 135). Put simply, the therapist must notice that the therapy interaction has gone awry. Recognition of conflict in the therapeutic relationship necessitates close attention to the interpersonal presentation of clients, particularly attention to subtle shifts in verbal and nonverbal behavior. Notably, this skill is probably more important with clients who respond to problems by withdrawing, rather than through overt confrontation. Although deceptively simple, conflicts in the relationship often “snare” therapists, such that they are contributing to the problem but remain unaware (Binder & Strupp, 1997; Knox & Hill, 2009; Safran & Muran, 2000), partially because clients often do not explicitly disclose negative reactions (Hill, Thompson, Cogar, & Denman, 1993; Hill, Thompson, & Corbett, 1992).

Once conflict is noticed, the therapist must possess the ability to tolerate conflict without withdrawing or making a counter-attack. For instance, when conflict is openly addressed often “therapists will respond with their own anger and defensiveness” (Safran

& Muran, 2000, p. 154; see also Hill et al., 2003). For example, if a client accuses a therapist of being ineffective (“I don’t know how you could possibly understand and can’t see how you can help”), the therapist may experience anger and want to blame the client (“I wonder if this has something to do with your relationship with your mother”). The therapist does not need to “avoid or to transcend” negative reactions altogether, but ruptures are more likely resolved when the therapist is able to recognize their emotions and maintain an objective and helpful perspective (Safran & Muran, 2000, p. 155; see also Hill & Knox, 2009). In short, tolerating interpersonal conflict involves self-awareness (e.g., “I’m feeling angry and threatened”), emotional understanding of self and others (e.g., “she’s probably feeling angry too”), and ability to manage one’s own emotional response (e.g., calming one’s feelings of anger).

The ability to self-reflect and manage internal reactions enables the therapist to maintain a therapeutic, or open, warm, empathic, comfortable stance (Binder & Strupp, 1997; Hill & Knox, 2009; Safran & Muran, 2000). Implied here is the notion that this therapeutic stance cannot be “faked” because genuinely hostile or blaming reactions couched in empathic or warm content are not hidden from the client and have a detrimental consequence for outcomes (Henry, Schacht, & Strupp, 1986, 1990). The therapists’ ability to maintain genuine warmth and empathy enables the interaction to move forward productively (e.g., the client explores their feelings and learns something new about themselves). In other words, the ability to accept, understand, and cope with internal reactions is important so that genuine therapist warmth and empathy can be expressed, negative reactions are not inadvertently communicated to the client, and the process can move forward.

Finally, ruptures in the alliance are negotiated through “relational processing” (Hill & Knox, 2009), or complex therapeutic skills (e.g., immediacy, metacommunication) that help the client understand and explore what is happening in the relationship (Binder & Strupp, 1997; Safran et al., 2002). Relational processing involves a therapist response (empathic and warm) that acknowledges or brings the problem that is occurring in the interaction between therapist and client into immediate focus, and successfully invites or engages the client in a dialogue about the current conflict. For instance, in the example above, the therapist might reflect that the client seems angry or disappointed, venture an empathic guess about what might be going on for the client, or openly invite the client to more specifically disclose their experience and feelings.

Overall, successful resolution of conflict is a complex skill set that requires the therapist to simultaneously attend to the client interpersonal content, process their own reactions, cope with internal reactions that aren’t therapeutic (e.g., hostility, neglect, blaming), respond in a therapeutic (e.g., warm, empathic, open) manner and engage in an immediate dialogue about the conflict. Therapists who do not have these abilities run the risk of “confound[ing] their interpersonal reactions to their patients and their clinical judgments” (Binder & Strupp, 1997, p. 124) and attributing blame to the client’s psychopathology or interpersonal problems. Notably, successful resolution of ruptures is a complex skill set that novice trainees are unlikely to possess. However, the most interpersonally talented trainees may be able to respond therapeutically even when it is challenging to do so. In short, successful resolution of alliance ruptures can essentially be boiled down to open, warm, empathic responsiveness during difficult moments and may be “at the heart of the change process” (Safran et al., 2002, p., 236).

Alliance Ruptures and Client Outcome. Existing research provides compelling support for the notion that alliance ruptures are fundamentally important to the therapy process and indicates that the manner in which these problems are addressed (or not) influences client outcome (Binder & Strupp, 1997; Couthino et al., 2011; Hill & Knox, 2009; Mohr, 1995; Safran & Muran, 2006; Safran et al., 2002). This seems particularly true for challenging clients who present for therapy with a poor prognosis, severe and chronic pathology, or pervasive interpersonal problems (Asay & Lambert, 2002; Binder & Strupp, 1997).

Extant evidence indicates that the strength of the therapy alliance changes over the course of therapy, and these changes often predict client outcomes (Safran et al., 2002). Fluctuations in the strength of the alliance over time are indicative of negative vs. positive interpersonal processes, and imply the presence of a rupture/repair process that can shape the course of therapy (Kivlighan & Schmitz, 1992; Kivlighan & Shaughnessy, 2000; Strauss et al., 2006). For example, in one study, optimal alliance development was dependent on client interpersonal problems. Specifically, in clients with adaptive relationships and interpersonal patterns, a strong stable alliance (average scores) was predictive of outcome. In contrast, clients with interpersonal problems benefited more from an alliance that increased over time (Piper, Boroto, Joyce, McCallum, & Azim, 1995). In other words, it may be that ruptures occur more frequently with clients who have a difficult interpersonal history, and successful resolution of ruptures allows the therapy to progress. This study demonstrates how positive changes in the alliance are more important with clients who are interpersonally challenging.

Therapist interpersonal responses can facilitate or impede progress with clients who are particularly challenging or are at risk for “treatment failure” (Suh, Strupp, & O’Malley, 1986). In a study intended to examine how therapists facilitate the interpersonal process, investigators divided clients on two dimensions: poor vs. good prognosis and poor vs. good outcome (Suh, Strupp, & O’Malley, 1986). They found “striking differences” in terms of how effective vs. less effective therapists changed their responses across the first three sessions to produce differential outcomes. Specifically, a small portion of clients ($n = 4$) with a *poor* prognosis achieved a *good* outcome. In these cases, increased expression of therapist warmth resulted in large increases in client participation (Suh, Strupp, & O’Malley, 1986). In contrast, a small portion of clients ($n = 2$) had a *good* prognosis but had *poor* outcomes. Sessions with these clients were characterized by high “negative therapist attitudes” (intimidating or threatening) which increased while warmth decreased. Essentially, the least effective therapists in this study were unable to effect change in clients with a good prognosis, whereas the most effective therapists were able to effect change in clients with a poor prognosis. This study illustrates that positive or negative therapist interpersonal responses can redirect the predicted course of therapy for clients in a negative or positive way.

Research continues to suggest that negative therapist responses to challenging or hostile client reactions have a detrimental consequence for client outcomes. For instance, a higher frequency of client hostility/blaming leads to a higher frequency of therapist counterhostility, a reciprocal pattern that has been related to poor alliances and client outcomes in several studies (Coady, 1991a; Henry, Schacht, & Strupp, 1986, 1990; Kiesler & Watkins, 1989; Klee, Abeles, & Muller, 1990; Navajits & Strupp, 1994;

Samstag et al., 2008; Tasca & McMullen, 1992; von der Lippe et al., 2008). These studies support the notion that when therapists do not recognize and remove themselves from negative interpersonal process “the relationship reaches an impasse, becomes stalemated, and aborts” (Strupp, 1998, p. 23; see also Coutinho et al., 2011). In addition to subtle therapist hostility and blaming, research suggests that therapists misuse specific techniques when presented with negative interpersonal process (Castonguay et al., 1996; Piper, Azim, Joyce, McCallum, Nixon, & Segal, 1991; Piper, Ogrodniczuk, Joyce, McCallum, Rosie, O’Kelly, & Steinberg, 1999).

Although several studies suggest that negative therapist reactions tend to be associated with poor client outcomes, only a few studies have directly examined whether therapists who demonstrate higher levels of interpersonal skills (positive therapist responses) in a challenging interpersonal situation produce superior outcomes. Specifically, data from a relatively large sample of therapists ($n = 25$) and clients ($n = 1,141$) in a naturalistic setting suggested that global ratings of therapist “facilitative interpersonal skills” (e.g., empathy, warmth, verbal fluency) were strongly related to client outcomes ($r = .47$; there were no differences among clients prior to treatment). Interpersonal skills in this study were rated on the basis of eight analogue “interactions” with videotaped actor clients specifically enacting interpersonally challenging clinical scenarios (Anderson et al., 2009). In an unpublished report using the same performance task, Anderson and colleagues (2001) similarly found that doctoral level student therapists with low facilitative skills were less effective with clients relative to students with high facilitative skills (as cited in Anderson et al., 2009).

Along similar lines, recent research has directly examined whether resolving alliance ruptures (positive therapist response) facilitates better client outcome. In one study intended to test the effectiveness of brief relational therapy (BRT; Safran & Muran, 2000), a treatment specifically designed to help therapists resolve alliance ruptures, Muran and colleagues (2005) found that clients receiving BRT therapy were less likely to drop out of therapy (20%) compared to short-term psychodynamic therapy (46%) or cognitive-behavioral therapy (37%); however there were no significant differences in terms of symptom changes (Muran, Safran, Samstag, & Winston, 2005). Consistent with this, research suggests that when therapists are given external feedback about problems in the alliance and encouraged to empathize and apologize, clients are less likely to deteriorate and outcomes are improved (Harmon et al., 2007).

Taken together, research suggests that the resolution of negative interpersonal processes that occur throughout the course of therapy partially determine the success of therapy (Safran et al., 2002). This evidence implies that therapists who get “snared” into negative interpersonal process are less successful in helping clients. There seem to be some therapists who work effectively with clients who are unlikely to make progress in therapy. Essentially, it may be that therapists (trainees) who can continue to demonstrate empathic and warm interpersonal stances “when the rubber meets the road” may become particularly talented therapists with training (Anderson et al., 2009; Binder & Strupp, 1997; Knox & Hill, 2009). Thus, testing the ability to tolerate negative reactions and maintain a therapeutic stance during interpersonal conflict seems particularly worthwhile for identifying individuals who are likely to become talented clinicians and evaluating interpersonal skills in trainees, but has rarely been examined.

Empirical Directions for Evaluating Interpersonal Skills

Taken together, clinical wisdom and research findings support the notion that interpersonal skills (e.g., empathy, warmth, and resolution of conflict) are fundamentally important to becoming a successful psychotherapist (Ackerman & Hilsenroth, 2001, 2003; Asay & Lambert, 2002; Beutler et al., 2004; Henry & Strupp, 1994; Lambert & Barley, 2002; Norcross, 2002; Teyber & McClure, 2000). Further, there is evidence to suggest that interpersonal skills form the basis for professional interactions beyond psychotherapy, such as clinical supervision and mentorship (Bordin, 1983; Ellis & Ladany, 1997; Gray, Ladany, Walker, & Ancis, 2001; Ladany, Friedlander, & Nelson, 2005; Ramos-Sanchez, Esnil, Goodwin, Riggs, Touster, Wright et al., 2002).

This is hardly revolutionary, as longstanding notions of the “ideal trainee” revolve around the belief that interpersonal skills are crucial to becoming an effective professional psychologist (APA, 2006b; APA Committee on Clinical Training, 1947; Johnson & Campbell, 2002, 2004; Korman, 1974; O’Donovan & Dyck, 2001; Peterson, 2003; Sakinofsky, 1979). Further, the recent competency movement in professional psychology includes relationship ability as a foundational competency that trainees should master over the course of training and across all professional functions (Assessment of Competency Benchmarks Work Group, 2007; Elman et al., 2005; Kaslow et al., 2007; NCSPP, 2007), yet highlights the need for students to be interpersonally skilled from the outset of training (Hatcher & Lassiter, 2007).

Thus, identifying and evaluating trainees on the basis of these interpersonal skills seems critical for both selection and evaluation over the course of graduate training. Nonetheless, the extent to which these skills can be measured reliably and meaningfully,

and the extent to which they predict clinical performance in trainees, is an issue that has received scant attention. To date, three assessment approaches relevant to evaluating trainee interpersonal skills have been examined over the last several decades for their potential use in selection and early evaluation procedures in professional psychology (and related) programs: 1) self-report, 2) cognitive performance tasks, and 3) complex interpersonal performance tasks.

Self-Report Measures

In the self-report approach to trainee evaluation, self-report measures intended to assess stable characteristics or personality traits are completed by trainees and then examined with regard to predictive validity. Several researchers have evaluated a variety of self-report inventories that are theoretically relevant to assessing relational ability. A number of self-report inventories, such as those intended to capture empathy (Hill, Roffman, Stahl, Friedman, Hummel, & Wallace, 2008; Stahl & Hill, 2008), emotional intelligence (Easton, Martin, & Wilson, 2008; Martin, Easton, Wilson, Takemoto, & Sullivan, 2004) and a wide array of personality traits (e.g., dominance, psychological mindedness) have been proposed as potentially useful for assessing interpersonal skills (Beutler et al., 2004).

Despite the notion that self-report measures are theoretically relevant and represent a potentially useful method for assessing interpersonal skills, research does not provide much support for this type of evaluation tool. For example, in a recent study of the effectiveness of “helping skills” training in a sample of undergraduates ($n = 85$), the investigators evaluated whether self-reported empathy and perfectionism predicted better interpersonal effectiveness in a simulated helping interaction with a peer. Although the

results indicated that trainees generally improved according to observer ratings of their “helping skills” after training, self-reported trait empathy and perfectionism were unrelated to performance in the helping interaction (Hill et al., 2008). In a similar study of undergraduates ($n = 125$), several self-report inventories (e.g., empathy) were unrelated to “client” perceptions of helpfulness in a simulated laboratory helping interaction (Stahl & Hill, 2008).

Overall, evidence contradicts the usefulness of self-report inventories for evaluating training interpersonal ability. This method of evaluation has been largely unrelated or inconsistently related to interpersonal performance assessed during interactions (Beutler et al., 2004; Dooley, 1975; Kelly & Fiske, 1951; Rappaport, Chinsky & Cowen, 1971). Consistent with this, professionals often overestimate their own abilities (Dooley, 1975; Gurman, 1977; Kurtz & Grummon, 1972) and therapist perceptions of their own skills are weakly related to client outcomes and client-rated perceptions of the same skills (Orlinsky et al., 2004). This evidence suggests that self-reported characteristics may or may not translate to performance in therapy-like contexts and situations. Furthermore self-report measures are unlikely to reflect the fundamental ability to perform well in interpersonally complex or conflictual situations.

Performance Tasks

In contrast to evidence regarding self report measures, there is a great deal of research to suggest that sampling specific desired behavior is the strongest indicator of the ability to adequately perform that behavior in the future (McHolland, Peterson, & Brown, 1987; Peterson, 2003; Schmidt & Hunter, 1998). Consistent with this notion, a handful of

different approaches to sampling relevant therapy-specific behaviors by means of performance-based tasks have been evaluated.

Cognitive Performance Tasks. In one common performance-based approach, cognitive abilities relevant to skillful interpersonal communication in therapy are sampled. For instance, therapist interpersonal skills have been inferred from tasks intended to assess the ability to identify emotions (Jackson, 1986), to analyze videotaped therapist behavior (Constanzo & Philpott, 1986; Loo, 1979), or to perform on a combination of tests intended to assess a cognitive skill set (e.g., interpersonal intelligence) relevant to interpersonal behavior (Constanzo & Philpott, 1986; Osipow & Walsh, 1973). In a prototypical study of this kind, undergraduate participants ($n = 40$) were asked to identify positive and negative therapy interactions and interpret interpersonal behaviors viewed on tape (Constanzo & Philpott, 1986). These cognitive abilities were moderately to strongly ($\beta = .38-.48$) predictive of interpersonal talent (e.g., warmth, empathy) demonstrated in an analogue helping scenario conducted in a classroom setting.

The findings of this representative study (Constanzo & Philpott, 1986) and other studies (Jackson, 1986; Osipow & Walsh, 1973) suggest that trainee performance on cognitive tasks is significantly predictive of supervisor ratings or interpersonal performance in benign analogue scenarios. However, there is a paucity of evidence to suggest that the ability to respond effectively to a simple cognitive task translates into performance in complex, clinical situations (i.e., in a therapy setting). Moreover, there is limited evidence to suggest that these cognitive tasks actually capture fundamentally important interpersonal skills. For instance, there is reason to believe that the ability to

identify or recognize warm or empathic behavior may not consistently translate into the ability to deliver a warm or empathic response, particularly in the context of a conflict. Indeed, evaluation of simple cognitive skills fails to capture the complex ability to respond effectively in challenging interpersonal situations.

Complex Interpersonal Performance Tasks. Although there is limited evidence to suggest that trainee performance on cognitive tasks is related to clinical outcomes, there is compelling evidence to indicate that performance in complex therapy-like performance tasks represents the most promising method for evaluating trainee interpersonal skills (APA, 2006a; Dooley, 1980; Kaslow et al., 2007). Whereas cognitive performance tasks assess simple skills, interpersonal performance tasks attempt to measure interpersonal skills in a context that approximates the complexities of the clinical encounter. Only a few interpersonal performance tasks exist (Anderson et al., 2009; Carkhuff, 1969; Goodman, 1972), most likely because it is challenging to create feasible performance tasks that adequately mirror the complexities of psychotherapy (i.e., are valid). Similarly, it has been difficult to operationalize complex interpersonal skills. To further complicate this issue, effectively measuring therapist interpersonal skills has not been at the forefront of investigation, with most contemporary efforts focused on improving treatment packages rather than practitioners. As a result, existing tools are largely outdated and inadequate.

For instance, to evaluate the effectiveness of a skills training program, Carkhuff (1969) designed a performance task (oral/written response to standard audiotaped/written vignettes) and rating scale intended to capture several interpersonal skills from the client-centered tradition (e.g., empathy, warmth, honesty). This measure generally seems to

represent a more effective approach for measuring relational ability because it directly assesses trainee responses to simulated (and at times challenging) clinical material and attends to several key interpersonal skills. Despite its promise, the Carkhuff (1969) rating scale has been heavily criticized by client-centered advocates (Lambert & DeJulio, 1977) and contemporary authors (Bohart & Greenberg, 1997b). For example, the scale is thought to misrepresent key interpersonal stances, which are boiled down to simple response sets (e.g., reflection of feeling equals empathy; Lambert & DeJulio, 1977; Lambert et al., 1978). Additionally, the Carkhuff (1969) rating procedure has not been linked with client outcome. Rather this measure has solely been utilized to measure the acquisition of skills rather than to assess how these skills are implemented in clinical scenarios (Lambert & DeJulio, 1977; Gormally & Hill, 1974).

A more promising, but still outdated performance-based tool was developed to assess “psychotherapeutic talent” in paraprofessionals (Goodman, 1972) and has garnered some attention and empirical support. Specifically, the Group Assessment of Interpersonal Traits (GAIT) was designed to select a talented group of nonprofessional counselors for the purposes of a community-based program for at-risk youth. The GAIT is a group performance task with a global rating scale based on Rogers’ (1957) client-centered conception of empathy, warmth, and congruence (Goodman, 1972). In this task, potential counselors are divided into small groups (6-7), prompted to respond to a personal question (intended to capture openness), and then take turns as “disclosers” and “understanders” during five-minute turns. “Understanders” are instructed to focus on emotions, provide their thoughts, or listen, but told not to give advice, make judgments, or ask questions. Participants and observers evaluate each group member on the basis of

7-items (e.g., understanding, open, quiet, warm, rigid) rated on 6-point Likert scale (Goodman, 1972).

The reliability and predictive validity of the GAIT for the selection of paraprofessional counselors has been evaluated in several studies to date (for a review see Dooley, 1980), the most recent study of which was published nearly 30 years ago. Only a few studies have evaluated whether GAIT-measured interpersonal skills predict client outcomes (Cahill, 1981; Chinsky & Rappaport, 1971; Dicken, Bryson, & Kass, 1977; Goodman, 1972; Rappaport, Chinsky & Cowen, 1971). For example, in one study, staff member ratings ($n = 3$) of global therapeutic talent (based on empathy, warmth, openness) were used to select paraprofessional counselors ($n = 99$; Goodman, 1972). GAIT ratings were generally consistent across raters, with acceptable staff inter-rater reliability for global talent ($r_m = .51$; $r_m = .54$ for empathy, warmth, and openness) and modest to moderate inter-rater agreement among staff and participant ratings ($r = .23 - .52$). In a sample of young males ($n = 88$), staff GAIT ratings of therapeutic talent had small, but nonsignificant associations with better self esteem ($r = .17$) and decreased aggression at school ($r = -.19$), and a significant modest association with improved adjustment ($r = .26$; Goodman, 1972).

Extant research generally supports the reliability of GAIT ratings (Dooley, 1980; Dooley, Lange, & Whiteley, 1979; Goodman, 1972) and there is preliminary evidence to suggest that the GAIT has weak to modest predictive validity in terms of clinical outcome (Cahill, 1981; Chinsky & Rappaport, 1971; Dicken, Bryson, & Kass, 1977; Goodman, 1972; Rappaport, Chinsky & Cowen, 1971). Although there is preliminary support for use of the GAIT method for selecting paraprofessional counselors and the GAIT does

well to closely approximate a live interaction as it would occur in therapy, the GAIT is limited in several important respects. For instance, the GAIT procedure is difficult to standardize, given that trainees respond to different individuals each time the GAIT procedure is administered. Furthermore, the GAIT rating scale is based on outdated definitions of interpersonal skills. For example, empathy is construed simply as the ability to understand whereas contemporary definitions emphasize the complex nature of this construct. Finally, the GAIT is based on performance in the context of a friendly (albeit anxiety provoking) interaction with a peer that does not capture the ability to respond to and resolve interpersonal conflict.

Recent evidence regarding the potential for interpersonal conflict to shape the course of therapy suggests that effectively responding to and resolving conflict is a routine part of the therapists' job (Binder & Strupp, 1997; Safran et al., 2002). As such, this interpersonal ability is important to include in tasks intended to capture the realities of therapy. Along these lines, the most promising interpersonal performance-based tool for identifying and evaluating professional trainees on the basis of their interpersonal skills was developed and utilized by Anderson and colleagues (2009) to predict therapist effectiveness in a large naturalistic, clinical sample.

The Facilitative Interpersonal Skills (FIS) task (Anderson, Patterson, & Weiss, 2006) is based on contemporary literature suggesting that a particularly promising way to evaluate the potential for clinical success may be to assess interpersonal effectiveness during an emotionally or interpersonally challenging situation. Thus, in the FIS task, therapists are asked to respond "as if" they are the therapist to standardized, videotaped clinical vignettes (based on real therapy clients), which are intended to simulate several

diverse, interpersonally challenging encounters. The accompanying rating scale (FIS scale; Anderson et al., 2006) is based on the most comprehensive, contemporary empirical literature available (Beutler et al., 2004; Orlinsky et al., 2004; Norcross, 2002).

The FIS performance task and rating scale have only been examined in two studies to date, one of which is an unpublished report. Ratings on the FIS scale were reliable across raters ($r > .70$) and were strongly related to client outcomes ($r = .47$) in one sample of mixed professional ($n = 19$) and trainee ($n = 6$) therapists and their clients ($n = 1,141$) in a naturalistic setting (Anderson et al., 2009). In an unpublished report that was exclusively focused on doctoral students (in psychology and non-helping disciplines), Anderson and colleagues (2001) found that the FIS performance task and rating scale were related to better client outcomes in brief therapy (7 sessions; as cited in Anderson et al., 2009). Taken together, these two studies provide preliminary evidence for the potential predictive validity of the FIS task and scale.

Notably, it is challenging to create a scenario that mirrors an actual clinical encounter, and any analogue task that seeks to approximate therapy with a real client, who is in distress, and needs help, will almost certainly fall short to a degree. For instance, the FIS performance task attempts to simulate challenging, realistic clinical material, but sacrifices some ecological validity, given that responding to a videotape is not the same as interacting with a live human. However, the FIS performance task is based on contemporary literature and, despite its limitations, was a strong predictor of client outcomes. Regardless of the challenges inherent in evaluation tools that rely on therapy-like simulations, complex interpersonal performance-based tasks are more likely to recreate the realities of psychotherapy. The FIS task in particular may provide an

accurate estimate of trainee interpersonal ability given that trainees must demonstrate their skills during a simulated interpersonal conflict.

In sum, a small handful of feasible and potentially useful evaluation tools and methodologies relevant to the measurement of interpersonal skills in trainees have been identified and sporadically evaluated; however, there are several limitations in the available research to date. First, trainee interpersonal skills have rarely been examined with regard to clinically relevant and practically important outcomes, such as effectiveness with therapy clients. Second, existing rating scales are largely outdated and do not adequately capture contemporary definitions of interpersonal skills. Finally, only one published study to date has incorporated evidence suggesting that the most promising therapists are interpersonally skilled and able to remain therapeutic (warm, empathic, open) during difficult interpersonal interactions (Anderson et al., 2009). Although the Anderson et al. (2009) study was not focused on trainees, interpersonal behavior samples taken during a task intended to simulate the complex realities of therapy seem to represent the most promising method for meaningfully evaluating fundamental trainee interpersonal skills that shape the course of therapy.

The Present Study

The purpose of this dissertation was to examine the reliability, discriminant validity, predictive validity, and incremental validity of the FIS performance task and rating scale. Toward this end, trainees enrolled in the UNLV clinical psychology doctoral program and the marriage and family therapy master's program were recruited to complete the FIS performance task (Anderson et al., 2006; Anderson et al., 2009). Trainee responses were

rated on the basis of the FIS rating scale. To determine whether ratings of trainee performance on the FIS task were related to clinical effectiveness, client data were obtained from the UNLV therapy training clinic and served as an estimate of therapist effectiveness. There were five specific aims.

The first aim was to determine whether FIS ratings of trainee interpersonal performance could be made reliably. Given the findings of previous research based on a wide range of scales tapping various interpersonal skills, I expected to find that the ratings in this dissertation would be reliable across raters with a minimal amount of training.

The second aim was to determine whether there were systematic differences in FIS performance based on stable therapist characteristics (e.g., age, gender, and training program type). There is some evidence to suggest that interpersonal performance on the FIS task improves with age (Anderson et al., 2009). Thus, I expected to find that performance would be positively associated with age. I did not expect to find systematic differences in interpersonal performance with regard to gender or training program type.

The third aim was to explore whether performance on the FIS task would capture a quality or set of qualities that differ from qualities easily captured via more traditional (GRE, GPA, number of practicum hours) and/or more efficient (brief self-report) evaluation methods. More specifically, this aim was geared toward exploring whether trainee interpersonal performance was related to academic performance/aptitude, time in training (i.e., clinical experience), and trainee self-assessment of interpersonal effectiveness with clients. With regard to academic performance/aptitude, there is scant

evidence to suggest how interpersonal performance is related to academic ability measures.

Although quantity of experience is often equated with better performance, few studies exist to determine whether trainees improve with training and experience. Studies that do exist are inconsistent, with some suggesting that interpersonal skills remain the same (e.g., Mallinckrodt & Nelson, 1991), some indicating they deteriorate (e.g., Henry et al., 1993), and still others that they improve (Crits-Cristoph et al., 2006; Hill, 2004; Hill et al., 2008; Hilsenroth et al., 2002; Truax & Carkhuff, 1967). I expected to find improvements in trainee empathic performance with more time in training, but did not expect to find changes in the other FIS performance areas.

The third aim was also focused on exploring whether a self-report measure of interpersonal performance was related to observer ratings of trainee interpersonal performance. Given the findings of previous research (Dooley, 1975; Gurman, 1977; Kurtz & Grummon, 1972; Orlinsky et al., 2004) I expected to find that self-report and observer measures of performance would be weakly related.

The fourth aim was to examine whether trainee interpersonal performance predicted therapist effectiveness (assessed via client outcome). Given previous research (Anderson et al., 2009; Goodman, 1972), I expected to find that better FIS performance would be associated with better clinical outcomes.

The fifth aim was to examine whether FIS performance predicted clinical performance above and beyond measures of academic aptitude/performance. Given previous research, I expected to find that academic aptitude and clinical performance

would be weakly related or unrelated (e.g., Bergin & Jasper, 1969), and that interpersonal performance would explain unique variance in trainee clinical success.

CHAPTER 3

METHODOLOGY

Participants

Eligible participants included any trainees who were enrolled in the clinical psychology doctoral program and marriage and family therapy (MFT) master's program and who had completed or would complete a future therapy practicum at the UNLV Center for Individual, Couple and Family Counseling (CICFC), a department sponsored campus-housed community mental health training clinic.

Recruitment

Three methods were used to recruit eligible trainees during a 6-month recruitment period (September, 2010 through March, 2011). First, an undergraduate research assistant (RA) visited graduate practicum classes and provided a brief (5-10 minutes) overview of the study purpose and procedures. Trainees were given a form on which to privately indicate whether they were interested in participating. Second, the RA invited all eligible trainees to participate via four separate emails that were addressed to each eligible individual and included the primary investigator's name and signature. Finally, I contacted all eligible trainees via email to provide an overview of the purpose of the study and study procedures.

A total of 38 clinical psychology students and 11 MFT students were eligible to participate and were actively recruited ($N = 49$). Of eligible participants, 38.8% ($n = 19$) completed the study materials. Three additional individuals agreed to participate (6.1%) but were contacted multiple times and failed to schedule an appointment. The remaining individuals ($n = 27$) refused participation either by declining to participate (32.6%) or via

passive non-response to email (22.5%). Recruitment was more successful with clinical psychology trainees, nearly half of whom participated in the current study ($n = 17$; 45% of initial recruitment sample). In contrast, two of eleven MFT trainees (18%) agreed to participate.

Demographics

The majority of participants were female (73.7%) Caucasians¹ (78.9%) with an average age of 27.7 years ($SD = 5.7$). As noted previously, the majority of the sample was enrolled in the clinical psychology program (89.5%). Participants represented a wide range of years in training (from 1 – 7) with most participants in the 3rd year of training (31.6%), followed by the 2nd year of training (21.1%).

With regard to level of education, 47.3% of the sample reported having already completed one Master's degree. Two individuals in the sample had attained two Master's degrees. Overall, participants had attained an A-B average in both undergraduate (M UGPA = 3.70, $SD = .17$) and graduate level education (M GGPA = 3.90, $SD = .21$) with academic programs reporting UGPAs ranging from 3.43 to 3.97 and GGPAs ranging from 3.10 to 4.00. Participant GRE total scores ranged from 1010 to 1480 ($M = 1254.74$, $SD = 130.95$) with higher GRE Quantitative scores ($M = 681.05$, $SD = 75.93$) than GRE Verbal scores ($M = 573.68$, $SD = 81.05$).

With regard to quantity of clinical experience, participants reported a wide range of overall clinical experience. Specifically, participants reported an average of 4.5 previous semesters ($SD = 4.2$) in therapy practicum training provided by their current graduate program, ranging from 0 to 18 semesters of supervised therapy experience. The majority

¹ To maintain participant confidentiality, ethnicity is not reported here given that individuals of ethnic backgrounds other than Caucasian are highly identifiable to program authority figures.

of participants (63.1%) identified their primary theoretical orientation as being represented by a combination of more than one theory including: Integrative, Eclectic, or an idiosyncratic listing of more than one primary theoretical orientation (e.g., Humanistic/Cognitive, Interpersonal/Humanistic).

Participant Therapy Client Caseloads

Therapy client files were eligible for inclusion in this dissertation if 1) the client was seen on an individual basis (vs. couple or family), 2) the client was over the age of 18, 3) the client attended at least three therapy sessions, and 4) the client provided consent for the use of their data in research on paperwork routinely collected at CICFC. A total of 18 trainees had begun to see therapy clients at CICFC by the end of the record review period. Therapists had an average therapy client caseload size of 10.22 ($SD = 4.29$) and a total of 184 therapy client files were located and reviewed for the purposes of this dissertation. Of these 184 therapy client files, a total of 123 files were not eligible for inclusion in the study because a) clients were under the age of 18 or received treatment as a couple/family ($n = 77$) or b) the clients did not provide consent for their information to be included in research ($n = 42$). A small portion of client files ($n = 4$) were not included in the study because the client attended fewer than 3 therapy sessions. Finally, 5 therapy client files could not be located. This left a total 56 therapy client files that were available and eligible for inclusion in the study or an average of 3.11 therapy client files ($SD = 1.88$) included in each therapist's caseload (range = 1 – 8; mode = 2).

Instruments

Measures in this dissertation were intended to assess 1) trainee demographic characteristics and FIS interpersonal performance as predictor variables, 2) therapist clinical effectiveness on the basis of CICFC client outcomes as the criterion variable, and 3) supplementary information regarding client levels of distress. Trainee measures were administered solely for the purposes of this dissertation, whereas the client measures are routinely administered and maintained at CICFC.

Trainee Measures

Trainee measures were intended to assess basic demographic information, academic aptitude/performance, quantity of training experiences, and interpersonal ability.

Therapist Characteristics. Demographic information (age, gender, ethnicity; see Appendix A) was obtained from participants via self-report. Additionally, participants were asked to provide their current year in training, type and years of any previous experiences relevant to helping professions, and their current theoretical orientation. Participants also provided consent for the RA to obtain official Graduate Record Examination (GRE) and Grade Point Average (GPA, both undergraduate and graduate) scores from program administrative staff.

Facilitative Interpersonal Skills (FIS) Performance Task. The FIS performance task (Anderson et al., 2006) was used to elicit interpersonal responses which were evaluated to obtain a global index of trainee interpersonal performance, as well as measures of specific relational abilities (e.g., empathy, warmth, emotional expression). The FIS performance task was developed to assess therapist interpersonal skills in response to depictions of a variety of interpersonally challenging clients. Specifically,

this task was designed to measure the ability to “perceive, understand, and communicate a wide range of interpersonal messages, as well as a person’s ability to persuade others with personal problems to apply suggested solutions to their problems” (Anderson et al., 2009, p. 759).

This task includes interpersonally challenging clients based on archival video segments of actual therapy sessions drawn from the Vanderbilt I and II research studies (Henry, Schacht et al., 1993; Henry, Strupp et al., 1993; Strupp & Hadley, 1979). The segments drawn from this archival research study for the purposes of the FIS task were chosen to reflect negative interpersonal process between therapist and client. Actors were asked to memorize transcripts from selected therapy sessions and to re-enact the interpersonal style of two overall patterns of interpersonal behavior: 1) “highly self-focused, negative, and self-effacing” and 2) “highly other-focused, friendly, but highly dependent” (Anderson et al., 2009, p. 759). Actors practiced prior to being videotaped and faced the camera to simulate a therapy-like situation. The task included one videotaped vignette (about 2 minutes) for each of seven different therapist-client dyads (see Appendix B for complete instructions and two sample transcripts).

Generally, each of the 7 vignettes is intended to simulate an alliance rupture “episode” (Anderson et al., 2006). The alliance ruptures reflected in the FIS vignettes capture a wide range of interpersonal styles that range from direct expressions of hostility toward the therapist to implicit and subtle interpersonal client patterns that threaten healthy development of the therapy alliance. For instance, the FIS performance task includes 1) a client who is “confrontational and angry” (“You can’t help me”), 2) a client who is “withdrawn, passive, and silent” (“I don’t know what to talk about”), 3) a client

who is “confused and yielding” (“only the therapists’ opinion matters”), and 4) a client who is “controlling and blaming” (“implies that others, including the therapist, are not worthy of him”; Anderson et al., 2009, p. 759). According to Anderson and colleagues (2006), the variability within the vignettes captures “interpersonal flexibility” or the ability to respond effectively to a wide range of clients with different interpersonal styles.

Facilitative Interpersonal Skills Observer Rating Scale (FIS-O). The FIS observer rating scale (FIS-O, Anderson et al., 2006) was used to evaluate the interpersonal effectiveness of trainee responses to the FIS performance task. The FIS-O is rooted in the empirical psychotherapy literature and captures those relational skills which seem to be the most clearly and consistently related to positive psychotherapy outcomes. Specifically, the FIS-O is an 8-item rating scale that captures therapist verbal fluency, emotional expressiveness, persuasiveness, hopefulness, warmth, empathic accuracy (attunement) and communication, alliance-bond capacity (collaboration), and alliance rupture/repair responsiveness (Anderson et al., 2006). The items are rated on a Likert scale ranging from 1 (*not characteristic*) to 5 (*extremely characteristic*). More specifically, ratings of 1 or 2 are intended to be reflective of skill deficiencies, a rating of 3 is intended to be neutral, and ratings of 4 or 5 are intended to be representative of a high skill level, or skill proficiency (Anderson et al., 2009).

Previous research provides preliminary support for the reliability and predictive validity of the FIS performance task and rating scale. For example, in one study, two independent licensed psychologists were able to attain adequate inter-rater reliability for all items ($r > .70$; Anderson et al., 2009). As was reviewed above, the FIS performance task and rating scale have demonstrated preliminary predictive validity in terms of client

outcomes. For example, ratings on the basis of the FIS performance task were moderately related to better clinical performance as assessed by client outcomes ($r = .47$) in one published report (Anderson et al., 2009) and one unpublished report (as cited in Anderson et al., 2009).

Ratings in this dissertation were based on the unpublished coding manual designed for the FIS task and the observer rating scale (Anderson et al., 2006). The original unpublished coding manual was revised slightly to 1) clarify construct definitions and 2) clarify and enhance rating anchors. For example, the FIS item description intended to capture persuasiveness included a reference to Jerome Frank's (1991) "believable myth" construct which is a technical reference to a concept included in a common factors theory. The definition of "believable myth" was added to the item description for raters who were unfamiliar with such technical references. Other items were revised to clarify questions that were raised by the team of raters throughout training meetings. Revisions were made with the intention to retain the meaning of the original FIS constructs. Generally, raters had the most conceptual difficulty with persuasiveness and emotional expression.

The FIS-O items for each of the seven FIS performance vignettes (for a total of 56 single items for each participant) were rated by four independent raters in this dissertation. For each individual rater, ratings for each of the eight FIS-O items were summed across the seven vignettes and ranged from 11 to 35 (with potential item scores ranging from 5 to 35). One grand total FIS performance score was calculated for each trainee by summing the item total scores and ranged from 135 to 257 (with potential total

scores ranging from 56 to 280) across the four raters. Ratings were condensed² across the four independent raters to obtain one single FIS total performance score and eight item scores for each participant (see Table 1 for condensed mean and standard deviation scores).

Table 1
FIS Performance Condensed Mean Total and Item Scores

FIS Observer Rating Scale	M	SD
FIS Total Scores	192.67	20.43
Verbal Fluency	23.28	3.23
Emotional Expression	25.34	2.72
Persuasiveness	25.62	2.04
Warmth	24.50	3.79
Hope	23.00	3.70
Empathy	23.09	3.28
Alliance Bond	26.24	3.81
Alliance Rupture Repair	21.60	3.85

Data from three unpublished research studies have suggested that FIS-O scores were internally consistent ($\alpha > .70$; Anderson et al., 2006) and FIS-O scores were internally consistent in this study sample ($\alpha = .90$). The data in this study indicated that FIS-O items were very highly correlated with FIS-O total scores ($r \geq .80$) with the exception of verbal fluency (see Appendix C). Verbal fluency was not significantly correlated with FIS-O total scores ($r = .04$, *ns*) or with any of the other FIS-O items. In fact, verbal fluency correlated with 6 of 7 items (emotional expression, warmth, hope, empathy, alliance bond, and alliance rupture repair) in a negative direction (α if item deleted = .95). In contrast, the other 7 FIS-O items were highly and significantly correlated with one

² Given that one of the primary aims of this dissertation was to evaluate the inter-rater reliability of the FIS performance task and associated rating scale, the inter-rater reliability results and justification for the use of condensed ratings are included in the results section.

another. Thus, results for FIS total performance scores will be reported. Item level analyses will also be reported if they diverge from findings using the global interpersonal performance score (FIS-O total scores)³.

Therapist Interpersonal Skills Self Report Inventory (TIS-SR). The TIS Self Report Inventory (TIS-SR) is a 30-item questionnaire (see Appendix D) that was developed for the purposes of this dissertation and was intended to parallel the FIS-O. Lists of 7-10 items were generated to capture each of the eight constructs described in the FIS-O manual. Approximately 3-4-items were included in the final measure to capture each FIS-O construct. Participants were instructed to rate each item on a 1 (*strongly disagree*) to 5 (*strongly agree*) Likert-type scale to indicate their beliefs about their ability to build and manage therapy relationships and communicate interpersonal messages in the therapy relationship. The mean total TIS-SR score was 121.84 ($SD = 6.73$) and participant scores ranged from 110 to 134 (with potential total scores ranging from 30 to 150). Internal consistency was marginally acceptable for the TIS-SR items ($\alpha = .68$).

Client Measures

Client measures were intended to assess and describe client initial level of distress and functioning and client level of distress and functioning at the end of therapy.

Outcome Questionnaire – 45 (OQ-45). The *Outcome Questionnaire – 45* (OQ-45; Lambert, Morton, Hatfield, Harmon, Hamilton, Reid et al., 2004) was used to assess and obtain a numerical value of client initial and final level of distress and functioning. These initial and final estimates of client distress levels were then used to calculate a client

³ FIS total scores were also re-calculated without scores on the verbal fluency scale and all analyses reported below were re-run. Removal of the verbal fluency scale score from FIS total scores did not change the results.

change score, which served as an estimate of how much each client changed over the course of therapy.

The OQ-45 is a widely used, 45-item self-report assessment of general client psychological distress and functioning. It was designed to track client progress in treatment settings. Clients are instructed to think about the past week and rate each item (e.g., “I have thoughts of ending my life,” “I feel lonely,” “I feel nervous”) on a 5-point Likert scale from 0 (*never*) to 4 (*almost always*). The OQ-45 yields a total score (ranging from 0 to 180) and three subscale scores (Symptom Distress, Interpersonal Relations, and Social Role). Higher total OQ-45 scores are indicative of more psychological distress. Factor analytic results suggest that the three OQ-45 scales are highly intercorrelated and may best represent a single dimension of psychological distress (Umphress, Lambert, Smart, Barlow, & Clouse, 1997). Thus, only OQ-45 total scores were included in the results reported below.

Research suggests that the OQ-45 has strong psychometric properties. The OQ-45 is internally consistent ($\alpha = .70 - .93$; Lambert et al., 2004; Vermeersch, Whipple, Lambert, Hawkins, Burchfield, & Okiishi, 2004) and has high test-retest reliability ($r = .84$ over a 3-week interval; Lambert, Burlingame, Umphress, Hansen, Vermeersch, Clouse et al., 1996). Research also supports the validity of the OQ-45 and suggests that it is significantly and positively correlated ($r = .50 - .85$; Lambert et al., 2004) with several widely used symptom measures, such as the Symptom Checklist – 90, the Beck Depression Inventory, and the State-Trait Anxiety Inventory. Furthermore, the OQ-45 is sensitive to client progress and changes in distress levels as a result of psychotherapy (Vermeersch, Lambert, & Burlingame, 2000; Vermeersch et al., 2004).

In this sample of therapy clients ($N = 56$), initial OQ-45 data were available for 46 clients. For the purposes of data collection and record review, initial OQ-45 scores were defined as any OQ-45 score that was obtained prior to or during the dates of treatment or therapy with the trainee therapist. Final OQ-45 data were available for 40 clients. For the purposes of data collection and record review, final OQ-45 scores were defined as the last available OQ-45 score obtained either while the client was engaged in treatment with the therapist or immediately after the ending of the therapy relationship with the therapist.

Both initial and final OQ-45 scores were available for a total of 36 clients that were seen by a total of 14 therapists (74%) in the sample. These clients had an average initial OQ-45 score of 71.83 ($SD = 20.60$) and an average final OQ-45 score of 63.56 ($SD = 29.00$). A portion of these clients (20.6%) had been seen previously by another therapist at CICFC. Clients were seen for an average of 19.72 sessions ($SD = 10.12$, range = 8 to 42) over the course of 6.28 months ($SD = 2.91$). These clients improved by an average of 8.28 OQ-45 points ($SD = 17.57$) but there was a large degree of variability with one client improving by 53 points and one client deteriorating by 21 points over the course of therapy.

The accuracy of the estimate of client change in therapy and thus, the accuracy of the estimate of therapist effectiveness in this study, was dependent upon the timing of the OQ-45 assessment. Therefore, the number of days between OQ-45 administration and the dates of therapy were examined. Although most of the initial OQ-45 scores were administered within a month of the first therapy session (76.5%), there was a great deal of variability with regard to when the initial OQ-45 was administered, ranging from 152 days prior to the first session for one client to 140 days after the first session for one

client. Most of the final OQ-45 scores obtained in this sample were administered either on the date of the final session or within 7 days of the final session (60.6%). However, there was also a great deal of variability with regard to the time between termination of therapy and final OQ-45 scores, with one score that was obtained 196 days before the final session.

Given the large degree of variability in the timing of OQ-45 administration, any initial or final OQ-45 score that was completed more than 6 weeks from the dates of therapy was removed from the sample (6 initial scores, 8 final scores). Additionally, there were 3 OQ-45 scores that were missing information regarding the timing of the OQ-45 assessment; these OQ-45 scores were also excluded.

Thus, out of a total of 56 therapy client files that were collected, 20 files were not eligible for inclusion because they did not include both initial and final OQ-45 scores (35.7%), and 17 files were removed due to variability in the timing of OQ-45 administration (30.4%). This left a total of 19 client files (33.9%) with both initial and final OQ-45 scores that were administered within 6 weeks of the dates of treatment (initial OQ-45 $M = 10.05$ days prior to first session, $SD = 10.95$ and final OQ-45 $M = 2.58$ days prior to final session, $SD = 7.05$) and were seen by a total of 10 therapists in the sample. These clients had an average initial OQ-45 score of 81.05 ($SD = 18.83$) and an average final OQ-45 score of 73.32 ($SD = 31.51$). A small portion of these clients ($n = 3$) had been seen by a previous therapist at CICFC. Clients with both initial and final OQ-45 scores were seen for an average of 13.50 sessions ($SD = 5.98$, range = 8 – 30) over the course of an average of 4.63 months ($SD = 1.88$, range = 2.10 to 7.97 months). This sample of clients improved on average by 7.74 points ($SD = 20.40$) but there was a large

degree of variability with regard to change in therapy with one client improving by 53 points on the OQ-45 and one client deteriorating by 21 points.

Termination Summary. Termination information in combination with a records review form was used to assess the number of therapy sessions and the dates of treatment for each client, including dates of treatment with therapists prior to the participant and therapists subsequent to the participant. The termination summary form is completed for each client upon the ending of therapy. The therapist completes this form and the form is signed by the therapist's supervisor. The termination summary form includes the total number of therapy sessions, treatment dates, a summary of the client's presenting problems and treatment goals, a summary of treatment progress and outcome, client final diagnosis, and recommendations for future treatment.

Procedure

Prospective participants were informed that the general purpose of the study was to examine how individuals respond to interpersonally challenging clients in a performance-based task and how these responses relate to clinical outcomes. Individuals who agreed to participate were asked to provide their contact information (email address and phone number) and scheduled an appointment with the RA who assigned each participant a random three-digit secret code. Participants also provided consent for the RA to access their therapy client files at CICFC and their academic records. Academic records (GPA and GRE scores) were accessed via departmental administrative staff.

Performance Task Administration

The RA administered all trainee materials during a scheduled appointment time. All study materials took less than one hour to complete. The RA escorted each participant into a private room, provided the instructions for the FIS performance task (“respond as if you are the therapist and leave the audiorecorder running”), turned on the audio-recorder, and left the room while the participant responded to the task. The RA was available to answer questions throughout material administration. Once the performance task was completed, the participants were asked to complete the self-report questionnaires. Upon successful completion of all study materials, each participant was given a \$15 Starbucks gift-card as a token of appreciation.

Performance Task Ratings

Ratings were completed in a series of steps. First, all recorded responses were transcribed to maximize the accuracy and reliability of ratings (Hill & Lambert, 2004).

Second, a team of four raters (advanced graduate students or recent post-doctoral professionals) blind to the identities of the therapists being rated were trained to complete the FIS-O response ratings. Raters included myself, two recent post-doctoral professionals (both from two different Clinical Psychology programs), and one Master’s level Counseling Psychology graduate student. Two of the raters trained in the UNLV Clinical Psychology program. The other two raters completed their training in different parts of the country.

Raters were provided with a coding manual for the FIS-O scale and were asked to study and become familiar with the manual. Sample recorded and transcribed responses to the FIS performance task reflective of a wide range of response types ($n = 5$) were

developed for training purposes. As a team, the raters completed one sample transcript and then independently rated four sample transcripts. Each sample transcript was completed prior to each team meeting. Team meetings included discussion of the constructs being rated and rater disagreements, as well as clarification and revision to the coding manual. The team met for a total of five, 1-1.5 hour long meetings prior to completing ratings for the participant data.

Finally, the raters were provided with four audio recordings and transcripts of participant responses to the FIS performance task at a time until ratings were completed for each participant. The rater team was instructed to primarily listen to each recording and to follow along with the transcript as needed (e.g., if they had trouble understanding the audio) as they completed each rating for each participant. All ratings were completed independently. Raters were instructed not to discuss ratings and were encouraged to consult the coding manual to resolve coding difficulties. In addition, raters were asked to provide comments about participant responses, the coding manual, and any difficulties they encountered while rating the responses. All ratings and descriptions of trainee interpersonal performance on the basis of the FIS task were finalized before any clinical outcome data was accessed. That is, raters remained blind to clinical outcome data.

Client Outcome Data

Client measures, including the OQ-45, are routinely collected as part of the procedures at CICFC and were obtained for the purposes of this dissertation. Client data was gathered as it became available throughout the course of the study. Specifically, complete client data was available for a portion of the trainees immediately upon their participation in the study ($n = 13$); whereas a portion of trainees were concurrently

completing their CICFC practicum placement ($n = 4$) or had not yet completed their CICFC practicum placement at the time of their participation in the study ($n = 2$). Thus, client data was collected for approximately 13 months and was completed in March, 2012. The RA collected all therapy client materials after signing an oath of confidentiality and being trained in client records review.

Lists of participants' CICFC individual adult therapy client numbers were obtained via several different methods. For several participants, a complete list of CICFC therapy client numbers was available via the CICFC archival database ($n = 8$). Unfortunately, the CICFC archival database was unavailable prior to 2009. Participants who completed their CICFC practicum placement prior to the implementation of the CICFC database ($n = 11$) were emailed by the RA and asked to locate and provide lists of their CICFC therapy client identification numbers. In addition, the RA searched every available CICFC file for participant therapist names. This search was conducted two times to maximize the number of therapy client files located for each therapist.

The therapy client data was obtained in several steps. First, the RA checked each relevant client record to determine whether he/she provided consent for his/her data to be accessed for the purposes of research. Only data for clients who provided permission for the use of their clinical information in research were included in this study. Second, the RA checked each relevant client record for other eligibility criteria (age, treatment modality, number of sessions > 2). Third, the RA photocopied relevant client data at CICFC and removed all therapist, clinical supervisor, and client identifying information. Specifically, the RA photocopied the initial and final OQ-45 score reports and termination summary as this information became available and completed a record

review form for each client. These materials were inaccessible to the rest of the research team until FIS ratings were finalized. These procedures were approved by the CICFC operating board and the UNLV institutional review board.

CHAPTER 4

RESULTS

Reliability of Ratings of Interpersonal Performance

The first aim of this dissertation was to determine whether ratings or judgments about the effectiveness of trainee interpersonal performance could be made reliably. To address this aim, I computed the inter-rater reliability of ratings of trainee facilitative interpersonal skills (based on the FIS-O) made by four independent raters. Specifically, the inter-rater reliability of FIS-O ratings (FIS-O total scores and 8 FIS-O item scores) was assessed using intraclass correlation coefficients (ICC; Shrout & Fleiss, 1979). ICCs were computed using a two-way mixed effects analysis of variance model, with raters as a random factor and agreement defined as absolute for 1) FIS-O total scores and 2) eight FIS-O items. Generally, values greater than .75 are considered “excellent,” values of .40-.75 are “fair-good,” and values below .40 are “poor” (Parkerson, Broadhead, & Tse, 1993). The results indicate that agreement for FIS-O total scores and for FIS-O items was excellent (all ICCs \geq .83, see Table 2).

Table 2
Inter-rater Reliability: FIS Ratings

FIS Observer Rating Scale	4 Raters
FIS Total Scores	.93
Verbal Fluency	.91
Emotional Expression	.84
Persuasiveness	.91
Warmth	.88
Hope	.90
Empathy	.93
Alliance Bond	.89
Alliance Rupture Repair	.83

I also computed ICCs for all possible combinations (4) of 3 raters and all possible combinations (6) of 2 raters using a two-way mixed effects analysis of variance model, with raters as a random factor and agreement defined as absolute for 1) FIS-O total scores and 2) the eight FIS-O items (see Table 3).

Table 3

Inter-rater Reliability Summary: FIS Ratings for 4-Raters, 3-Rater Combinations, and 2-Rater Pairs

FIS Observer Rating Scale	4 Raters	3 Raters ^a	2 Raters ^b
FIS Total Scores	.93	.82 - .92	.68 - .85
Verbal Fluency	.91	.75 - .88	.62 - .81
Emotional Expression	.84	.60 - .79	.37 - .79
Persuasiveness	.91	.51 - .92	.20 - .80
Warmth	.88	.83 - .86	.73 - .83
Hope	.90	.82 - .89	.71 - .91
Empathy	.93	.85 - .93	.74 - .91
Alliance Bond	.89	.79 - .87	.69 - .80
Alliance Rupture Repair	.83	.74 - .87	.54 - .86

^a ICC ranges are based on all four possible rater combinations.

^b ICC ranges are based on all six possible rater combinations.

Overall, for the four combinations of 3 raters, ICCs ranged from fair-good to excellent ($ICCs \geq .51$; see Table 4). Specifically, results indicate that agreement for FIS-O total performance scores was excellent for all 3-rater combinations. At the item level, ICCs ranged from fair-good to excellent, with most of the ICCs for item scores falling in the excellent range. In fact, for one 3-rater combination, the ICC value was slightly improved when compared to the ICCs computed for all raters for persuasiveness (.91 to .92) and for alliance rupture repair (.83 to .87). This 3-rater combination attained ICCs that all fall in the excellent range.

Table 4
Inter-rater Reliability: FIS Ratings, 3-Rater Combinations

FIS Observer Rating Scale	1x2x3	1x2x4	1x3x4	2x3x4
FIS Total Scores	.92	.84	.84	.82
Verbal Fluency	.88	.75	.82	.80
Emotional Expression	.76	.73	.79	.60
Persuasiveness	.92	.65	.61	.51
Warmth	.85	.83	.83	.86
Hope	.86	.89	.82	.84
Empathy	.93	.90	.86	.85
Alliance Bond	.87	.84	.79	.83
Alliance Rupture Repair	.87	.74	.74	.75

Overall, for the six 2-rater pairs, ICCs were more variable and ranged from excellent to poor (see Table 5). ICCs for FIS-O total performance scores fell in the excellent range (ICCs $\geq .77$) with the exception of one rater pair, for which the ICC fell in the fair-good range (ICC = .68). At the item level, the majority of ICC values fell in the fair-good to excellent range. The exceptions to this included ratings for emotional expression and persuasiveness.

Table 5
Inter-rater Reliability: FIS Ratings, 2-Rater Pairs

FIS Observer Rating Scale	1x2	1x3	1x4	2x3	2x4	3x4
FIS Total Scores	.85	.77	.78	.81	.68	.79
Verbal Fluency	.62	.81	.70	.74	.70	.72
Emotional Expression	.60	.69	.79	.37	.53	.61
Persuasiveness	.80	.60	.49	.65	.20	.41
Warmth	.78	.73	.74	.81	.79	.83
Hope	.91	.71	.78	.75	.82	.78
Empathy	.91	.80	.84	.84	.79	.74
Alliance Bond	.80	.69	.74	.77	.80	.72
Alliance Rupture Repair	.86	.71	.56	.77	.54	.72

In sum, analyses indicated that a high level of inter-rater agreement ($ICC > .75$) was attained with four independent raters on FIS-O total performance scores and item scores. As would be expected, inter-rater agreement decreased slightly for three raters; however, ICC values still fell in the excellent range ($ICC > .75$) for FIS-O total scores. In addition, three rater combinations were able to attain excellent inter-rater agreement for most of the items. Finally, inter-rater agreement was more variable for two rater pairs. The results presented below are based on mean FIS-O scores calculated by averaging the FIS-O ratings made by the four independent raters.

Interpersonal Performance and Stable Trainee Characteristics

The second aim of this dissertation was to explore whether trainee interpersonal performance (as assessed by the FIS task) was associated with stable therapist characteristics (age, gender, and training program type).

To examine the relationship between trainee age and interpersonal performance, I calculated a Pearson correlation. Contrary to expectations, there was no significant association between overall performance on the FIS task (FIS-O total scores) and trainee age ($r = -.08, p = .749$).

To assess the relation between overall interpersonal performance on the FIS task (FIS-O total scores) and gender, I calculated an independent samples t test. Females ($M = 194.79, SD = 21.93, n = 14$) were not significantly different from males ($M = 186.74, SD = 16.02, n = 5$) in terms of overall interpersonal performance ($t(17) = .75, p = .466$, Cohen's $d = .42$).

To assess the relation between overall performance on the FIS task (FIS-O total scores) and training program type, I computed an independent samples t test. The MFT trainees ($M = 200.75$, $SD = 24.04$, $n = 2$) did not perform significantly better than the clinical psychology trainees ($M = 191.72$, $SD = 20.60$, $n = 17$) with regard to FIS-O total scores ($t(17) = -.58$, $p = .569$, Cohen's $d = .40$).

In sum, results suggested that FIS performance in this sample was not significantly associated with stable trainee characteristics. Specifically, FIS performance did not seem to be associated with therapist age, gender, or graduate program type. For therapist gender and graduate program type, the differences were not statistically significant even though the effect sizes were moderate.

Interpersonal Performance and Traditional Evaluation Methods

The third aim was to explore whether FIS performance captured a quality or set of qualities that differed from qualities captured via traditional evaluation methods (GRE, GPA, number of practicum hours) and/or more efficient evaluation methods (brief self-report). This aim was addressed using a series of analyses intended to evaluate the relationship between FIS performance and academic performance/aptitude, quantity of experience, and scores on a brief self-report measure of interpersonal effectiveness.

To assess the relation between overall interpersonal performance on the FIS task and academic performance/aptitude, I computed Pearson correlations between 1) undergraduate Grade Point Average (UGPA) and FIS-O scores, 2) graduate GPA (GGPA) and FIS-O scores, and 3) GRE verbal (GRE-V), quantitative (GRE-Q), and total

(GRE-T) scores and FIS-O scores. None of the correlations were statistically significant (see Table 6)⁴.

Table 6
Academic Performance/Aptitude and Interpersonal Performance

FIS-O Rating Scale	UGPA (n = 13)	GGPA (n = 19)	GRE-V (n = 19)	GRE-Q (n = 19)	GRE-T (n = 19)
FIS Total Scores	-.24	.17	-.03	-.16	-.12
Verbal Fluency	.36	.05	.42 ^t	.28	.42 ^t
Emotional Expression	-.10	.36	.00	-.27	-.16
Persuasiveness	.06	.07	-.04	.03	.00
Warmth	-.27	.24	-.05	-.24	-.17
Hope	-.28	.40 ^t	-.16	-.28	-.26
Empathy	-.28	.02	-.12	-.25	-.22
Alliance Bond	-.43	.20	-.14	-.14	-.16
Alliance Rupture Repair	-.32	-.26	-.08	-.08	-.10

** $p < .01$. * $p < .05$. ^t $p < .10$.

The relationship between quantity of experience and trainee interpersonal performance was also examined. To explore this relationship, I calculated Pearson correlations between scores on the FIS-O and time (years) in training⁵. Results indicated that the association between overall interpersonal performance on the FIS task (FIS-O total scores) and time in training approached significance (see Table 7). At the item level, better performance on alliance rupture repair and empathy were significantly related to increased time in training. There was also an association between scores on the emotional expression item and time in training that approached significance.

Given the qualitative differences between the MFT and clinical psychology graduate training programs and the small number of MFT trainees ($n = 2$) who had the same

⁴ Although statistical trends are identified in Table 6, these statistical trends are not referenced in the text given the high Type I error rate.

⁵ The analyses that follow were also conducted for practical experience as measured by quantity of clinical training (number of semesters in practicum training). Given the very strong association between time in training and quantity of clinical training ($r = .91, p < .001$) the correlations between these two different measures of the quantity of experience and FIS performance were synonymous. For ease of presentation, the results for quantity of clinical training are omitted.

number of years of training, this analysis was repeated for the clinical psychology participants ($n = 17$). For the clinical psychology trainees, results indicated a statistically significant association between interpersonal performance and time in training (see Table 7). Specifically, year in training was significantly related to better overall interpersonal performance (FIS-O total scores), and higher scores on warmth, empathy, and alliance rupture repair. Year in training was also moderately related to higher scores on the emotional expression and alliance bond items at the trend level.

Table 7
Year in Training and Interpersonal Performance

FIS-O Rating Scale	Entire Sample	Clinical Psychology
FIS Total Scores	.46 ^t	.54*
Verbal Fluency	-.11	-.12
Emotional Express	.40 ^t	.47 ^t
Persuasiveness	.35	.38
Warmth	.37	.50*
Hope	.18	.25
Empathy	.48*	.60*
Alliance Bond	.41 ^t	.48 ^t
Rupture Repair	.68**	.72**

** $p < .01$. * $p < .05$. ^t $p < .10$.

In addition to traditional evaluation methods employed in graduate training programs, I explored the relationship between FIS performance and scores on a brief measure intended to capture trainees' self-assessment of interpersonal effectiveness. Toward this end, I calculated Pearson correlations between scores on the FIS-O and scores on the TIS self-report inventory (TIS-SR; see Appendix E). Results indicated that there was no significant relationship between overall performance on the FIS task (FIS-O total scores) and total scores on the TIS-SR ($r = .21, p = .388$). Generally these two methods of

assessment seemed largely unrelated (i.e., out of 81 correlations, only 3 were statistically significant).

In sum, interpersonal performance had no significant relationship with traditional and more efficient methods of evaluation, with the exception of time in training. FIS performance did not have a significant association with academic performance/aptitude or trainee self-perceptions about interpersonally effectiveness with clients. In contrast, interpersonal performance (particularly the ability to effectively address conflict in the therapy relationship) did tend to improve as quantity of time in training increased.

Interpersonal Performance and Clinical Effectiveness

The fourth and fifth aims of this dissertation were to assess the predictive (postdictive) and incremental validity of trainee interpersonal performance. To assess the predictive validity of trainee interpersonal performance I computed an average caseload change score for each therapist based on initial and final OQ-45 scores. First, I obtained an OQ-45 change score for each individual therapy client by subtracting the final OQ-45 score from the initial OQ-45 score so that higher caseload change scores would be indicative of client improvement and lower caseload change scores would be indicative of less change or client deterioration. Second, the change scores were averaged across therapy clients to obtain one average caseload change score per participant (see Table 8).

Unfortunately, average caseload change scores were available for only a portion of the sample ($n = 10$) and were calculated on the basis of only a portion of trainees' CICFC caseloads (see Table 8). In fact, for four therapists, only one change score was available so that the "average caseload change score" is actually based on one score rather than a

caseload change score averaged across a number of clients. This was due to a variety of factors, such as eligibility criteria for inclusion in the sample, but of the 56 therapy client files that were eligible for inclusion in the study, there was a substantial portion of those files that were missing either 1) initial OQ-45 scores, 2) final therapy OQ-45 scores, or 3) both initial- and final-therapy OQ-45 scores. In addition, a number of OQ-45 scores (18) were eliminated from the data analysis given the large variability in the timing of initial and/or final OQ-45 administration which called into question their usefulness for providing an accurate estimate of client change. Thus, the results should be interpreted with extreme caution given the small sample size and the substantial portion of data that are missing for each therapist.

Table 8
Therapy Client OQ-45 Change Scores and Percentage of Caseload Missing

ID	Client OQ-45 Change Scores	Average Change	# of Change Scores	Total # on Caseload	% Missing
1			0	0	0.00
2			0	10	100.00
3	-21, -10, 0	-10.33	3	11	72.73
4	13, 53	33.00	2	7	71.43
5			0	7	100.00
6	14	14.00	1	7	85.71
7	-9	-9.00	1	24	95.83
8	9, 12	10.50	2	13	84.61
9			0	6	100.00
10	10, 13	11.50	2	11	81.82
11	2, 40	21.00	2	12	83.33
12	37	37.00	1	9	88.89
13	-21, -20, 14, 17	-2.50	4	10	60.00
14	-6	-6.00	1	6	83.33
15			0	10	100.00
16			0	11	100.00
17			0	9	100.00
18			0	15	100.00
19			0	6	100.00

Note. Participants are in order from poorest to best overall performance on the FIS Task.

To determine whether therapist interpersonal performance predicted clinical effectiveness, I computed Pearson correlations between interpersonal performance on the FIS task (FIS-O scores) and average caseload change. Results suggested that there was no significant correlation between interpersonal performance and average caseload change (see Table 9).

Table 9
Interpersonal Performance and Average Caseload Change

FIS-O Rating Scale	Pearson r (n = 10)
FIS Total Scores	-.11
Verbal Fluency	-.16
Emotional Expression	-.23
Persuasiveness	.27
Warmth	-.22
Hope	-.04
Empathy	-.10
Alliance Bond	-.15
Rupture Repair	.04

** $p < .01$. * $p < .05$. ^t $p < .10$.

In addition, I examined the relationship between trainee self-assessment of interpersonal effectiveness and average caseload change. Specifically, I computed Pearson correlations between total scores on the TIS-SR (TIS-SR total scores and scale scores) and client change. There was no significant relationship between scores on the TIS-SR and average caseload change (see Table 10).

Table 10
*Therapist Interpersonal Skills Self Report
 and Average Caseload Change*

TIS-Self Report Inventory	Pearson r (n = 10)
TIS-SR Total Scores	-.04
Verbal Fluency	-.13
Emotional Expression	-.27
Persuasiveness	-.11
Warmth	-.36
Hope	.22
Empathy	-.36
Alliance Bond	.18
Rupture Repair	.35

** $p < .01$. * $p < .05$. ^t $p < .10$.

Given that the above results contradicted previous empirical research on the relationship between performance on the FIS task and client outcome, I explored whether other variables could better account for variation in therapist effectiveness in this sample. Toward this end, I examined the association between therapy-specific variables (client initial levels of distress, treatment length) and average caseload change. Although none of the correlations were statistically significant, this may be due to small sample size and correspondingly low power (see Table 11).

Similarly, I assessed the relationship between traditional therapist variables (age, time in training, quantity of practicum semesters) and average caseload change via Pearson correlations. None of the correlations were statistically significant (see Table 11).

Table 11
*Therapy Specific and Traditional Therapist
 Variables and Average Caseload Change*

Variable	Pearson r (n = 10)
Therapy Specific Variables	
Average Initial OQ Scores	-.42
Average Number of Sessions ^a	-.55
Average Days in Therapy	-.24
Traditional Therapist Variables	
Age ^a	.56
Year in Training	.12
Semesters in Practicum	-.07

** $p < .01$. * $p < .05$. ^t $p < .10$.

^a Data was missing for one participant (n = 9).

The final aim in this dissertation was to determine whether interpersonal performance as measured by the FIS task would predict (postdict) therapist effectiveness above and beyond measures of academic performance/aptitude. However, given that the FIS performance task was unrelated to average caseload change, I simply examined the associations between academic aptitude/performance (GPA and GRE scores) and average caseload change via Pearson correlations. None of the correlations were statistically significant (see Table 12). In addition, I computed a partial correlation between FIS performance and average caseload change after controlling for the effects of academic ability and aptitude (partial $r = .13$, $p = .868$), which was small and nonsignificant.

Table 12
*Academic Performance/Aptitude and
Average Caseload Change*

Academic Measure	Pearson r (n = 10)
UGPA ^a	.26
GGPA	-.35
GRE-Verbal	.17
GRE-Quantitative	.15
GRE-Total	.23

** $p < .01$. * $p < .05$. ^t $p < .10$.

^aUGPA scores were missing for 2 participants (n = 8).

In sum, performance on the FIS task was unrelated to therapist effectiveness as measured by average caseload change. A number of other variables were examined to determine whether they could better account for variation in therapist effectiveness in this sample; none of these variables, however, were significantly correlated with average caseload change. Specifically, a self-report measure of trainee facilitative interpersonal skills, initial levels of client distress, number of therapy sessions, therapist age, years in training, number of semesters in practicum training, and measures of academic performance/aptitude were unrelated to average caseload change. Although none of the correlations were statistically significant, this may be due to shortcomings of the data measuring client change and the small sample size and correspondingly low power.

CHAPTER 5

DISCUSSION

This dissertation was designed to explore the reliability, discriminant validity, predictive validity, and incremental validity of the FIS performance task and rating scale. In this section, the primary findings of this dissertation, its limitations, and its implications for practice and research are presented.

Primary Findings

Reliability of Ratings of Interpersonal Performance

The first primary finding of this dissertation was that overall performance on the FIS task could be rated reliably. Specifically, global performance on the FIS task was rated reliably by four independent raters and was also rated reliably by all combinations of three raters. As would be expected, there was more variability in ICCs for two rater pairs. Nonetheless, all ICCs for two rater pairs fell in the excellent range with the exception of one rater pair which attained an ICC in the fair to good range. This finding replicates previous research which suggested that ratings of performance on the FIS task can be made reliably (Anderson et al., 2009) and is consistent with a more general (albeit small) body of literature which indicates that interpersonal skills can be evaluated by multiple raters in a consistent manner (Dooley, 1980; Dooley, Lange, & Whiteley, 1979; Goodman, 1972). This finding departs from previous research in three ways.

First, whereas previous empirical research involving the FIS performance task focused on evaluating the interpersonal performance of a sample that included individuals ranging from post-graduate to highly seasoned licensed professionals, this

dissertation focused on determining whether interpersonal performance on the FIS task could be rated reliably in a sample of trainees enrolled in professional graduate training programs. The results in this study suggest that trainee FIS performance can be evaluated reliably.

Second, compared to previous research that utilized two independent licensed psychologists as raters, this study involved a diverse group of raters. Specifically, the raters in this study were at different points with regard to their level of training (two very recent post-docs, one advanced Ph.D. student, and one master's level trainee). This suggests that it is feasible to train individuals with various amounts of experience to reliably evaluate responses to the FIS task within a relatively brief training period (5-6 hours).

Third, although the results of this dissertation confirmed the reliability of FIS-O total scores, results were more variable with regard to ratings of specific interpersonal performance areas, particularly for the two rater pairs. Specifically, previous empirical research involving the FIS task reported consistently acceptable inter-rater reliability ($r > .70$; Anderson et al., 2009) at the item level for two independent raters. Estimates of inter-rater reliability for two rater pairs at the item level in this sample varied from poor to excellent. Whereas a good level of agreement was attained across rater pairs for several of the specific performance areas (including verbal fluency, warmth, hope, empathy, alliance bond), there was more disagreement among some rater pairs for ratings of emotional expression, persuasiveness, and alliance rupture repair. Differences between this dissertation and previous research could be due to differences in the participants (trainees vs. seasoned professionals), the raters (licensed psychologists vs.

early professionals), the number of rater pairs (i.e., because there were various combinations of rater pairs in this study, more variability in agreement is not surprising), or the use of different statistics (intra-class correlation coefficient vs. Pearson correlation).

Taken together, the reliability results of this dissertation add another piece of evidence to a growing body of literature that suggests reliably assessing interpersonal performance (via performance on the FIS task) is both possible and practical (Anderson et al., 2009; Dooley, 1980; Dooley, Lange, & Whiteley, 1979; Goodman, 1972). This finding stands in stark contrast to the relatively little attention paid to assessment of trainee interpersonal skills during the selection process and in the evaluation methods utilized thereafter by graduate training programs. The neglect of assessing trainee interpersonal skills has been attributed in part to challenges in defining and reliably measuring interpersonal competencies despite their relevance to successful clinical performance (APA, 2006a; Lichtenberg et al., 2007).

Although the quantitative findings in this study seem to provide a clear answer to questions about whether interpersonal skills can be evaluated reliably in a sample of trainees (yes, they can, at least as defined by the FIS task), this study also may provide some insight into the difficulties training programs face when attempting to evaluate interpersonal performance. Indeed, observations made during the coding process suggested that disagreements (though infrequent) were emotionally charged and powerful and seemed to confirm that there certainly are times when evaluating interpersonal performance can feel complex, subjective, and difficult to define or adequately capture. For example, in one case, a rater expressed very positive feelings toward a sample

responder (stated that she would very much like to have her as her therapist) whereas another rater demonstrated a very negative reaction to the same responder (stated that he found it very difficult to evaluate the responses favorably because of a high degree of anxiety inherent in the responses). Furthermore, these reactions substantially influenced their ratings (inflating scores for the first rater and diminishing them for the second). Essentially, these raters were influenced by their emotionally charged, idiosyncratic reactions to this person, while the other two raters remained unaffected and largely agreed in their ratings of the same person.

This trend continued throughout data coding and was observed in an overall high level of agreement *most* of the time but some substantial disagreements between raters on the quality of *a few* participant responses. These disagreements were observed both quantitatively (via large disagreements between FIS total scores) and in rater comments about each participant. For example, one participants' interpersonal performance was ranked very highly by one rater, ranked very poorly by another rater, and ranked as middle of the road by the other two raters (M FIS total = 184.25/rater totals = 159, 179, 185, 214). From one rater's perspective, the interpersonal responses were characterized by "authority," "confidence," and "great cadence and melody." In contrast, another rater described the overall interpersonal feeling in the responses as "above it all," "condescending," and "belittling." Given that the two other raters agreed about this participant's overall performance, it seems possible that these descriptors were representative of opposing personal reactions that these raters experienced, which were then reflected in their quantitative rankings.

The underlying cause of rater conflict for only a few select participants remains unclear. For example, these disagreements could have been related to the individual participants themselves (i.e., certain participants responded in a way that is interpersonally complex or sends a double message) or to the individual raters (i.e., interaction effect or poor fit between specific raters and specific participants). Despite the fact that questions remain about the source of the disagreements in this study, it seems possible that the emotionally charged personal reactions demonstrated by raters in this study might parallel difficulties trainers face when attempting to evaluate the interpersonal effectiveness of their trainees (e.g., attempting to maintain an objective evaluation despite a strong personal reaction, and/or facing a respected colleague who completely disagrees). Overall, the complexities and anomalies observed in this study may inform our understanding of the lack of formal assessment/evaluation focused on trainee interpersonal effectiveness and deserve further exploration.

In sum, consistent with previous research, the FIS performance task did seem to provide a standardized method with which to reliably evaluate trainee interpersonal performance. Inter-rater reliability was particularly strong for total interpersonal performance scores. Furthermore, the findings of this study went above and beyond previous research to suggest that FIS ratings can be made reliably in a sample of trainees and with a sample of raters with various levels of experience. In contrast, agreement was somewhat more variable with regard to the specific performance areas (e.g., FIS-O items, particularly emotional expression, persuasiveness, and alliance rupture repair). Finally, observations made regarding rating disagreements may provide some insight into challenges faced by training programs when evaluating interpersonal performance.

Overall, the findings in this study demonstrate that dimensions of interpersonal skills can be measured reliably and therefore provide a promising step in the direction of determining the potential utility of the FIS.

Interpersonal Performance and Traditional Evaluation Methods

The second primary finding of this dissertation was that performance on the FIS task did seem to measure a quality or a set of qualities that diverged from qualities captured via more efficient methods (e.g., brief self-report) and/or via traditional evaluation methods (e.g., GRE, GPA). Interpersonal performance as captured by the FIS task was not significantly related to a self-report measure of interpersonal effectiveness or academic aptitude/performance. Better performance on the FIS task was related to more years of training (particularly for the clinical psychology trainees). In other words, the FIS performance task captured something that wasn't easily observable or captured via more traditional methods of assessment and something that improved with more time in training. Each of these key findings will be discussed.

Trainee Self-Evaluation. The results of this study suggested that trainee self-evaluation was unrelated to observer ratings of interpersonal performance. In other words, higher ratings on a self-report measure of facilitative interpersonal skills were not significantly correlated with higher observer ratings of those same skills. Although this result is limited by several factors, most importantly by the use of a self-report measure that has no empirical support for its reliability or validity (TIS-SR), this finding is consistent with previous research that suggests that self-report and observer measures of therapist qualities are often unrelated (Bohart et al., 2002; Dooley, 1975; Gurman, 1977; Kurtz & Grummon, 1972; Orlinsky et al., 2004). Specifically, in a review of the

relationship between empathy and client outcome, Bohart and colleagues (2002) found that therapist self-reported empathic ability and observer ratings of empathy were consistently unrelated. The authors hypothesized that this may be due to the tendency for therapists to overestimate their own abilities (Bohart et al., 2002). Research has also suggested that therapist self-report is often unrelated to client outcome or client ratings of therapist skills (Orlinsky et al., 2004). Thus, the lack of a significant correlation found in this study is consistent with previous research.

The reason for the lack of association between therapist self report and observer ratings has yet to be determined both in this study and in the larger research literature. What does seem clear across several different research studies is that these two different methods of assessment seem to capture different qualities.

Academic Ability. Although better academic aptitude and performance as measured by undergraduate GPA and Graduate Record Exam scores are the most heavily weighted factors at admission for most programs, and graduate GPA is an important determinant of progression through graduate school, the results of this study suggested that these measures are unrelated to interpersonal performance, at least as measured by the FIS-O. The reason for the lack of relationship between FIS performance and academic performance/aptitude remains unclear. It is possible that a larger sample size would have produced significant correlations. However, it is also possible that academic ability measures and interpersonal performance are genuinely unrelated.

The notion that academic and interpersonal abilities are unrelated is consistent with the “germ theory myth” (e.g., “exposure to coursework leads to catching the skill bug”; Ladany, 2007, p. 392) of training and education in the mental health field. In other

words, the results in the current study provide another bit of evidence to suggest that reliance on coursework performance (UGPA and GGPA) or measures of academic aptitude (GRE) to evaluate or to predict interpersonal ability (as it is defined by the FIS-O) is inappropriate.

Quantity of Experience. The results of this study suggested that better interpersonal performance was associated with increased time in training. Specifically, there was a significant relationship between overall interpersonal performance on the FIS task and time in training for the clinical psychology trainees ($r = .54$). In addition, all but one of the correlations between the specific interpersonal performance areas and time in training were in the expected direction (i.e., positive and $\geq .25$ with the exception of verbal fluency). In particular, clinical psychology trainees with more training were observed to demonstrate significantly more warmth ($r = .50$), more empathy ($r = .60$), and a better ability to address and repair conflict in the therapy relationship ($r = .72$). Although the correlations were smaller and there were fewer statistically significant relationships, results that included MFT trainees were similar.

Given that training programs expressly aim to build clinically relevant interpersonal skills, these results support both the validity of the FIS task as a measure of those skills and the general impact of training programs. With regard to the validity of the FIS task, this finding makes sense and provides some support that the FIS task actually captures something that is relevant to what trainees in professional training programs are learning. Along these lines, two of the three specific performance areas that demonstrated the strongest relationship with quantity of training (empathy and ability to repair conflict in the alliance) are areas that could reasonably be expected to improve with more

experience. For example, responding to conflict in the therapy relationship is a complex ability that involves several different skills (e.g., attending to interpersonal process, emotion regulation) and empathy is often a primary target of clinical training.

Unfortunately, there is very little empirical research that can speak directly and specifically to the link between the FIS task and quantity of clinical training/experience. Research on the relationship between these two constructs as they are measured more broadly (e.g., clinical skills instead of interpersonal skills; years of experience instead of training) in the literature is inconsistent (e.g., Crits-Cristoph et al., 2006; Henry et al., 1993; Hilsenroth et al., 2002; Mallinckrodt & Nelson, 1991; Hill, 2004; Hill et al., 2008; Truax & Carkhuff, 1967) and uses a wide range of methodologies that differ from the methodology used in this study. For example, changes in therapist skills are generally not directly assessed (Crits-Cristoph et al., 2006; Hilsenroth et al., 2002; Mallinckrodt & Nelson, 1991) or studies include samples of experienced therapists (e.g., Henry et al., 1993). Another body of literature does suggest that clinical skills improve with training; however this link is based on measurement of clinical micro-skills (e.g., the use of specific clinical behaviors, such as a “reflection of feeling” or “affirmation,” Hill, 2004; Hill et al., 2008; Truax & Carkhuff, 1967) and discrete training methods rather than a more general therapeutic interpersonal stance and graduate training overall.

Only one other empirical study has examined the validity of the FIS task (Anderson et al., 2009). This study was focused on a more experienced sample of therapists rather than trainees and as a result, did not directly evaluate the relationship between quantity of training and ratings of interpersonal performance. Notably, there was a moderate correlation between scores on the FIS task and age ($r = .45$), about which the authors

hypothesized that “age serves as an indicator of the accumulation of clinical experience needed to master the interpersonal qualities inherent in FIS” (Anderson et al., 2009, p. 764). Although somewhat different from the findings in this study that quantity of training, rather than the passage of time alone (age) was related to better performance on the task, these findings are consistent with one another and suggest that whatever construct the FIS captures does improve with quantity of clinical experience collected over the years (Anderson et al., 2009) *or* quantity of time in clinically oriented graduate programs.

In sum, this study provides preliminary evidence that the FIS task captures what it intends to capture and may do so in a manner that goes beyond traditional methods of assessment. Specifically, consistent with previous research (Bohart et al., 2002; Orlinsky et al., 2004), the results indicate that the FIS task measured something different than what is easily captured by a self-report measure. Performance on the FIS task also was unrelated to measures of academic aptitude or performance, providing evidence to suggest that these traditional academic measures fail to capture the important interpersonal skills that the FIS task intends to measure. In contrast, FIS performance was positively related to time in training, supporting the validity of the FIS performance task (i.e., our training programs *are* intended to improve performance). Overall, the FIS task seemed to measure a set of qualities that increased with time in training, as differentiated from academic aptitude/performance and trainee self-assessment, and it may very well be a useful tool for evaluating interpersonal skills throughout graduate training.

Interpersonal Performance and Therapist Effectiveness

Third and finally, the FIS task was tested with regard to its predictive and incremental validity. The correlations between performance on the FIS task and therapist effectiveness (as assessed by average caseload change) were nonsignificant and small. Results based on a self-report measure of interpersonal effectiveness were similar to findings with the FIS performance task. In addition, the relationship between measures of academic ability and therapist effectiveness were examined. They also were nonsignificant and modest. In this sample, client initial level of distress, number of therapy sessions, and therapist age seemed to demonstrate stronger associations with average caseload change scores than any of the measures of therapist characteristics or skills, but these results were still nonsignificant and could be due to random error.

Together, these findings are inconsistent with the intuitive notion that trainees who are interpersonally talented (based on observer ratings and/or self-report) are able to develop better therapy relationships with their clients and thus facilitate more client improvement. These findings are also inconsistent with previous research indicating that performance on the FIS task was related to higher levels of clinical effectiveness. More specifically, in one published report, performance on the FIS task was strongly related to client outcomes ($r = .47$; Anderson et al., 2009). In one unpublished report exclusively focused on trainees, Anderson and colleagues (2001) found that FIS performance was related to better client outcomes in brief therapy (as cited in Anderson et al., 2009).

Although performance on the FIS task failed to predict therapist effectiveness in this sample of trainees, the results involving client outcome data are highly limited and are inconclusive for three primary reasons. First, the power to detect statistical significance

in this study was limited by the very small number of trainees who were included in this analysis ($n = 10$) and the very small number of therapy clients. Indeed, the previous study that sought to examine the link between therapist interpersonal performance and client outcome involved data from over 1,000 therapy client files treated by 25 therapists (Anderson et al., 2009).

Second, the differences in the client populations alone could account for the different results. For example, the client population served by CICFC is known for severe distress levels, comorbidity, and personality pathology. As a result, clients who seek services at the center are often treated over the course of several years by multiple trainee therapists. Ideally, change scores are computed by taking a measure at the outset and termination of treatment; however, in this sample it was merely the beginning and ending dates of treatment with each trainee therapist (i.e., at the beginning and ending of practicum semesters). Essentially, this undermines the validity of the average caseload change scores as estimates of therapeutic effectiveness because it calculates them based on arbitrary dates that aren't related to symptom manifestation or amelioration.

Finally, it seems likely that the quality of the data that was used to estimate "therapist effectiveness" was highly questionable. The quality of the client outcome data was highly problematic because of a large number of client cases that were not eligible for inclusion in the study or because of data that was missing from the client files. Specifically, 90% of desired client data was either ineligible (67%), unavailable (i.e., the entire file was missing; 3%) or missing some portion of data (20%). As a result, "therapist effectiveness" was approximated for only a subset of the participants and was based on only a small percentage of each trainee's caseload, a modal number of 1-2

clients per participant. Overall, it seems highly problematic to make global judgments about “therapist effectiveness” based on such a small amount of information.

In addition to a large portion of missing therapy client data across all trainees, the missing data were also not evenly distributed with regard to scores on the FIS task. Specifically, rather than random participants having missing therapy client data, individuals at the top and bottom of the distribution of FIS scores were disproportionately impacted. Specifically, the five most interpersonally skilled participants and the two least interpersonally skilled participants based on FIS total scores were excluded from the analysis due to missing therapy client outcome data. The fact that data were missing for these particular therapists limited the range of FIS total scores and consequently the ability to detect a relationship if one in fact existed.

In sum, FIS performance was not related to “therapist effectiveness” but there are substantial reasons to doubt the validity of these findings. Thus, we should consider them inconclusive.

Limitations

This study was exploratory and had notable limitations. First, there were only 19 participants from two graduate training programs in this study. In addition, only two of these participants were from the MFT program. Although the size of this sample was comparable in size to the sample of therapists utilized by Anderson et al. (2009), as well as the larger literature on therapist effects, several of the analyses, particularly analyses intended to evaluate the validity of the FIS task, may have been limited by the small number of participants in the study. Despite the fact that some of the results regarding

the relationship between the FIS task and other forms of evaluation were theoretically coherent and meaningful, there was a high experiment-wise Type I error rate in this study and thus results could be attributable to random error.

Second, as noted previously, the number of complete therapy client files that were available and eligible for inclusion in this study highly limited the findings involving the predictive validity of the FIS task. It did not seem reasonable to believe that “average” caseload change scores could accurately depict “therapist effectiveness” given the small number of clients who were included in these scores. In addition, almost half of the already small sample ($n = 9$) was entirely missing any complete therapy client files and had to be excluded from the analysis.

Third, the definition of therapist effectiveness limited the study, both practically and theoretically. This limitation both involves the use of client outcome to define therapist effectiveness and the use of OQ-45 to define client outcome. Therapist effectiveness is a difficult construct to measure because client outcome is not solely dependent on the therapist, but is a dynamic interaction that involves the therapist, the client, the clinical supervisor, as well as environmental and contextual factors that cannot be easily measured. Although improvements in client symptom distress do ultimately seem to be the most practically important outcome when evaluating the effectiveness of professional and trainee therapists, reducing therapist effectiveness down to a client change score is a rather one-dimensional way to examine a process that is multidimensional and complicated.

Along similar lines, the use of the OQ-45 to define client change in this study boils another very complicated construct down to a simple number that may not actually

capture all the ways in which clients improve or change over the course of therapy. Although the OQ-45 is one of the most empirically validated and reliable measures of changes in client symptom distress over the course of therapy (Vermeersch, Lambert, & Burlingame, 2000; Vermeersch et al., 2004), the OQ-45 and any similar measure of symptom distress cannot account for the number of variables that may impact and influence the process of change in therapy. Essentially, the OQ-45 may not capture all the information that seems crucial to determining whether or not a client has improved as a result of therapy or as a result of their work with a particular therapist. It seems likely that this dissertation was limited by reliance on unidimensional measures to define realistically complex processes.

Implications for Practice

Despite limitations, the findings in this study do inform evaluation and assessment practices in graduate training programs, particularly given the recent competency movement which calls on all graduate training programs to evaluate all the qualities necessary to producing competent professionals (APA, 2006a; Hatcher & Lassiter, 2007; Kaslow et al., 2007; Nelson, 2007; Rubin et al., 2007) and longstanding concerns regarding current evaluation procedures (APA, 2006a; Beutler, 1995; Bickman, 1999; Kaslow et al., 2007; Ladany, 2007; McHolland, Peterson, & Brown, 1987). For example, current evaluation practices in training programs have been criticized for relying on quantity of practicum hours and graduate coursework to measure trainee professional competence (Bickman, 1999; Ladany, 2007). In fact, one researcher hypothesized that a large percentage of trainees who graduate are incompetent therapists partially due to the

invalid and ineffective evaluation and gatekeeping practices currently utilized in professional graduate training programs (Ladany, 2007).

The lack of formal evaluation methods geared toward measuring qualities that might contribute to clinical success is often attributed to the difficulties inherent in measuring such ambiguous and complex constructs. The results of this dissertation contribute to a small body of literature that suggests that it is indeed feasible to measure the interpersonal effectiveness of trainees in a manner that is consistent across different raters using a sample of relevant behaviors generated during a performance-based task that attempts to mirror the realities of the clinical encounter. In short, this study provides evidence that assessment of interpersonal ability, at least using the FIS performance task, could feasibly be done by advanced graduate students during the admission process or by trainers over the course of training.

The findings of this dissertation also bear directly on questions about whether current evaluation practices utilized at admission and beyond in graduate training programs adequately capture personal qualities that seem like they would be related to professional success. Specifically, the results of this dissertation lend some credence to the criticism that measures of academic ability utilized at admission (GRE, undergraduate GPA) and measures of academic performance utilized throughout graduate training (graduate GPA) have little to do with interpersonal effectiveness. In short, the results of this dissertation lend support to the assertion that “exposure to coursework” or ability to perform in the classroom does not lead to “catching the skill bug” (Ladany, 2007, p. 392). The implication of this finding is that traditional assessment methods used at admission

(GRE) and throughout graduate training (GPA) are not comprehensive enough for graduate training programs geared toward producing competent professionals.

On the other hand, the findings of this dissertation also provide some evidence to suggest that time in training is associated with improved interpersonal skills despite critics who have argued that programs tend to rely upon a “practice makes perfect” myth of professional psychology training (i.e., tendency to infer clinical capability based on the quantity of clinical experience; Beutler, 1995, 1997; Bickman, 1999; Ladany, 2007). In other words, the present findings suggested that time in training, at least in two training programs at one university, is associated with better interpersonal skills in *some* trainees, particularly their ability to express warmth and empathy, and their ability to directly address relationship conflict. Interestingly, these particular performance areas represent the skills in the empirical literature that are most strongly linked with better therapy outcomes.

The implications of this finding are two-fold. First, a measure of interpersonal performance could be used not only to evaluate trainees in terms of their interpersonal competence, but could also be used to evaluate the quality of training programs (i.e., is this training program having any influence on the interpersonal ability of its trainees?) or particular clinical supervisors, or particular practicum training sites. Second, the FIS performance task could feasibly be used to identify standardized interpersonal benchmarks in training programs and could help to identify those trainees who are not responding to training in an expected manner or who could simply benefit from targeted formative feedback. This identification process has the potential to help trainers provide targeted feedback that enhances trainee performance and informs plans for trainee

development or remediation. Of course, such an effort would be complicated by all of the difficulties inherent in high-stakes testing, such as “teaching to the test” and possibly increased pressure on both students and programs to cheat.

In short, enhanced training and evaluation procedures focused on clinically relevant trainee behaviors could ultimately have positive effects on client care. Although this study was exploratory and does not provide direct evidence of the link between performance on the FIS task and therapist effectiveness, the results of this study do provide evidence to suggest that the FIS performance task could represent a feasible way to provide valuable information about relevant skills prior to admission to graduate school and throughout graduate training.

Implications for Research

The results of this dissertation also have clear implications for research. This exploratory study was geared toward examining whether the FIS performance task is a reliable and valid measure of trainee interpersonal ability; several of the findings suggest that the FIS task is a measure worthy of future attention. It seems clear from the results of this study, as well as previous research (Anderson et al., 2009), that the FIS task is feasible to administer, can be rated reliably, and seems to capture a set of qualities that diverge from qualities captured by traditional evaluation methods. However, important questions remain about whether these findings generalize to different groups of trainees and raters. More research is needed before the FIS performance task can reasonably be used as an evaluation tool in graduate training programs.

It also remains unclear whether performance on the FIS performance task is meaningfully related to important training outcomes. Along these lines, it seems quite important for future research to determine whether performance on the FIS task and other assessment tools geared toward measuring personal qualities in trainees can predict various practically important outcomes. For example, in this study, client outcome was defined as a primary indicator of interpersonal effectiveness. However, there are several different ways in which interpersonal ability might be important during professional training programs. For example, interpersonal ability is likely to impact the development of relationships with professors, peers and mentors, as well as play a role in securing strong practicum and internship placements. Thus, research examining the relationship between FIS performance and various markers of success in graduate training seems warranted.

Finally, it seems crucial for future research to continue to examine individual therapist effects in determining therapy outcome. The more that research can identify which therapist variables contribute to success, the more equipped training programs will be to define and measure whether their trainees have what it takes to be clinically effective, are progressing as would be expected, and will be clinically competent professionals. Future research that explores individual therapist effects may serve to clarify which evaluation tools and methods would be best for training programs to use as gatekeepers of the profession.

APPENDIX A: Demographic Questionnaire

BACKGROUND INFORMATION

Please provide the following information about yourself:

1. Age: 0
2. Gender (Please check one): ☐ Male ☐ Female
3. With which of the following groups do you most identify?
☐ Caucasian or European American ☐ African American
☐ Hispanic-American/Latino(a)/Chicano(a) ☐ Asian American
☐ American Indian/Native American ☐ Pacific Islander American
☐ Other (please specify):
4. What type of graduate training program are you currently enrolled in? (Please check one):
☐ Clinical Psychology ☐ Marriage and Family Therapy
5. How long have you been enrolled in your graduate program: (Please check one)
☐ I am in my 1st year of training (12 months or less)
☐ I am in my 2nd year of training (13-24 months)
☐ I am in my 3rd year of training (25-36 months)
☐ I am in my 4th year of training (37-48 months)
☐ I am in my 5th year of training (49-60 months)
☐ I am in my 6th year of training or above (61 months or more)
6. What is your current theoretical orientation? If you do not yet know your theoretical orientation, please provide your current clinical supervisor's theoretical orientation. If you do not yet have a clinical supervisor, please check "not applicable." (please check one):
☐ Behavioral ☐ Integrative ☐ Systems
☐ Biological ☐ Interpersonal ☐ Other

- ☐ Cognitive Behavior ☐ Humanistic/Existential please specify: _____
☐ Eclectic ☐ Psychodynamic

7. How many semesters (fall, spring, summer) of therapy practicum training have you completed as a part of your current graduate training program? (please include the current semester): _____

8. Have you had any other previous training or practical experiences (not including your current training or sanctioned program activities) relevant to becoming a psychotherapist or counselor in mental health?

☐ No ☐ Yes

If you checked “yes” to question 6 above, please provide more detailed information about your relevant experiences.

Formal Training Experiences: *Formalized training experience* includes any time spent in a formal training program (e.g., intended end result is a degree or credential) relevant to the mental health profession (e.g., counseling psychology, social work) prior to enrollment in your current training program. If applicable, please list and provide information about any prior training in the space provided below.

Program type	Degree type	Degree completed?	Time in training	# of semesters of supervised therapy practicum?	Average # of therapy hours/week

Practical Experiences: *Practical experience* includes any time spent engaging in clinical activities that could be considered therapeutic for a client (e.g., psychotherapy, counseling, answering calls on a hotline) that were/are not sanctioned as a part of your current training program or that were not included under “formalized training experiences” above. If applicable, please list and provide information about these experiences in the space provided below.

Agency Type	Describe Duties	Were you supervised?	# years/mos at position	Average # of hours/week

9. Please list the most advanced degree you have received and the program type (e.g., Master of Arts in Clinical Psychology): _____

APPENDIX B: Facilitative Interpersonal Skills Task Sample Vignettes

This program will give you a chance to practice responding to a variety of therapeutic situations. The task should take you about 30 minutes and in that time you will be asked to give 7 open ended verbal responses. You will be asked to role-play, the helper, or therapist, in these situations. Brief videos will be used to encourage you to immerse yourself in the role of the therapist. To help you do this, most of the time you will be looking directly at the client, and your job is to pretend that you are the therapist or helper. There will be at least one point where the actual therapist says something to the client. You can expect the actual therapist to make maybe one intervention or so per clip. On the one hand, you may feel comfortable with what the therapist actually says, but on the other hand, you may feel that the actual therapist moves the conversation in a direction that you might not go yourself. Most of the videos last between one and two minutes and the longest is 5 minutes. Each of these videos were drawn from actual transcripts of psychotherapy interactions.

To help you get into the role, here are a few things to keep in mind. Each segment is from the third session of psychotherapy. That's usually enough time for client and therapist to be fully comfortable with each other. But as you will see that's not always the case. You'll get a very small amount of background information too, which may help you get into the role. Sometimes the video segment will be from the middle of the conversation and it may be a little difficult for you to understand the context. That's to be expected, so just do your best. It will be up to you to find out what they are discussing, what they are thinking, what they are feeling, and what each person wants to gain from the conversation. When the video clip stops, that means it's your turn to talk. The client has paused and is waiting for you to say something to be helpful. You can just go ahead and speak to the screen just as if you were in a real conversation. Try to speak as you actually would in the situation and say what you believe would be most helpful to the client at that point. Your response will be audio recorded for later analysis. Again, when the video stops, it will be your turn to talk.

1: JOHN

Background:

This is John and he'll be the first person that you'll be seeing today. He's depressed because his relationship with his fiancé ended several months ago and he still feels like he can't get past it. He is meeting you for his third session and he is pretty happy working with you so-far. However, he is really not sure how therapy is supposed to work even though he is feeling better after his first two visits with you. Remember when the video stops that means it's your turn to talk. Go ahead and talk just as if you are in a real conversation with John. Best of luck.

Therapy Clip:

John: What do you do after we talk? Do you write up notes? Or is it, or do you just sort of keep a running tally of where we are going since it's all on tape, or what?

Therapist: *No, I write up a short note. So what are you wondering, what I write or say or something?*

John: *Well, uh, I guess if it was written up as, as like a case study in a clinic, I just wonder what sort of jargon would be applied to me. Um, but I guess more generally, how you view my progress, and knowing what you know about how to progress, how you think it might progress further.*

Therapist: *Well...*

John: *But I don't know if you can answer that without suggesting, you know uh, prophecy or something like that [laughs]*

Therapist: *laughs*

4: SUZIE

Background:

This is Suzie. She is in her mid-thirties and she's depressed about being a lonely heart and other things. She yearns for excitement and thinks her life is often boring. Suzie has seen you before but she's frustrated because things still haven't worked out. When the tape stops I want you say something to Suzie as if you were talking to her. Say something that you would say naturally to someone like her. Again when the tape stops it will be your turn to talk.

Therapy Clip:

Suzie: *[in an agitated tone] Its not just upset with you, it's just that I keep asking for directions or guidelines or something, and I just don't feel like I ever get anything.*

Therapist: *Maybe somehow people try to respond to you or...*

Suzie: *[interrupts] No... [shaking head]*

Therapist: *...or be for you and it's like you feel like there's...*

Suzie: *[interrupts] No...[shaking head]*

Therapist: *...nothing anyone can say...*

Therapist: *[talking over one another] ...or do...*

Suzie: *[talking over one another] I don't think that's right...*

Therapist: *...that you can...*

Suzie: *...because I think people reach out, that people come to me, and I always have to reach out to them and help them...*

Therapist: *Mhhmmhh. But...*

Suzie: *I mean, I just...*

Therapist: *You just...*

Suzie: *Really, I just don't think, I just get more confused and more frustrated...*

Therapist: *Mhmmmmm.*

Suzie: *...and you know [big sigh, rolls eyes], I don't think that there's anything you can do to help and I don't know what I can do to help myself because...[shaking head]*

APPENDIX C: FIS-O Scale Correlation Matrix

Scale	FIS Total	Verbal Fluency	Emotional Expression	Persuasion	Warmth	Hope	Empathy	Bond
FIS Total								
Verbal F.	.04							
Emotion E.	.84**	-.08						
Persuasion	.83**	.26	.56*					
Warmth	.87**	-.26	.87**	.55*				
Hope	.86**	-.26	.82*	.62**	.84**			
Empathy	.97**	-.06	.75**	.79**	.85**	.80**		
Bond	.95**	-.13	.75**	.75**	.88**	.84**	.95**	
Rupture	.80**	-.03	.47*	.71**	.57*	.54*	.84**	.77**

** $p > .01$. * $p > .05$. ^t $p < .10$.

APPENDIX D: Therapist Interpersonal Skills Self-Report Inventory

THERAPIST INTERPERSONAL SKILLS INVENTORY

Directions: The following statements concern how effective you generally believe you are at building and managing therapy relationships as well as communicating interpersonal messages in the therapy relationship. Please read each statement and consider the extent to which it reflects your current or future ability as a therapist. Respond by indicating the degree to which you agree or disagree with each statement and write the number that corresponds with the rating scale in the space provided. If you have not yet been involved in a therapeutic or counseling relationship, please rate the degree to which you believe you will be able to effectively manage therapy relationships in the future.

1 Disagree Strongly	2	3 Neutral/ Mixed	4	5 Agree Strongly
_____	1.	When I am with a therapy client, I am emotionally expressive.		
_____	2.	I feel compassionate when I am with my therapy clients.		
_____	3.	I struggle to verbally communicate in therapy.		
_____	4.	I am comfortable addressing conflicts in the therapy relationship.		
_____	5.	I am able to help my clients feel hopeful about taking specific actions in therapy.		
_____	6.	My interventions effectively communicate that it is important for us to work together in therapy.		
_____	7.	I am emotionally engaged with my therapy clients.		
_____	8.	I am effective at convincing my therapy clients about my rationale for treatment, even when their views differ from my own.		
_____	9.	I clearly communicate my belief that therapy clients have the potential for change.		
_____	10.	I feel confident when I am speaking with my clients.		

1 Disagree Strongly	2	3 Neutral/ Mixed	4	5 Agree Strongly
_____ 11.	I directly discuss the interpersonal conflicts that I have with my therapy clients.			
_____ 12.	I am effective at focusing on my clients, even when I have a strong emotional reaction.			
_____ 13.	I actively work to establish a partnership with my therapy clients.			
_____ 14.	I have been told that I am sometimes condescending toward my therapy clients.			
_____ 15.	My clients and I are a team in therapy.			
_____ 16.	I frequently steer the therapy session to meet my own agenda.			
_____ 17.	I confidently explain my treatment rationale to my therapy clients.			
_____ 18.	I have a hard time persuading clients to follow my lead.			
_____ 19.	I often become irritated with my therapy clients.			
_____ 20.	I understand the deeper meanings of my therapy clients' thoughts, emotions, and experiences.			
_____ 21.	I feel anxious when I verbally communicate with my clients.			
_____ 22.	When I am with a therapy client, I keep my emotions to myself.			
_____ 23.	I am able to communicate my perspective in a convincing manner.			
_____ 24.	My therapy clients often tell me that I have not understood them.			
_____ 25.	My verbal communication in therapy is easy to follow.			
_____ 26.	I am able to help my clients feel capable of moving toward change.			

1	2	3	4	5
Disagree Strongly		Neutral/ Mixed		Agree Strongly

- _____27. I am attuned to the thoughts, feelings, and experiences of my therapy clients.
- _____28. I communicate caring concern for my clients.
- _____29. I recognize relational conflict as it occurs in my therapy relationships.
- _____30. It is easy for me to put my own concerns aside in therapy so that I can focus on my client.

APPENDIX E: Correlations: FIS Observer Scale (FIS-O) and TIS Self-Report (TIS-SR)

FIS Performance Task and TIS-Self Report Inventory: Pearson Correlations

FIS Performance	TIS Self Report Measure								
	TIS-SR Total	VF	EE	P	W	H	E	AB	RR
FIS-O Total	.21	.24	-.05	.16	.43 ^t	.18	-.08	.24	-.28
Verbal Fluency	-.08	.58**	-.24	-.31	-.29	-.36	-.01	-.21	.14
Emotional Expression	.26	.20	.09	.19	.36	.29	-.11	.04	-.08
Persuasiveness	.28	.42 ^t	-.25	.22	.27	.16	-.22	.49*	-.18
Warmth	.19	.07	.11	.12	.41^t	.30	-.03	.14	-.24
Hope	.06	.02	-.05	.21	.40 ^t	.27	-.21	.17	-.43 ^t
Empathy	.25	.19	-.02	.25	.45 ^t	.28	-.08	.31	-.32
Alliance Bond	.16	.17	-.09	.17	.43 ^t	.23	-.07	.26	-.36
Rupture Repair	.24	-.00	.08	.17	.53*	-.04	.13	.34	-.14

** $p > .01$. * $p > .05$. ^t $p < .10$.

Note. VF = Verbal Fluency, EE = Emotional Expression, P = Persuasiveness, W = Warmth, H = Hope, E = Empathy, AB = Alliance Bond, RR = Rupture Repair.

REFERENCES

- Aamodt, M.G., Bryan, D.A., & Whitcomb, A.J. (1993). Predicting performance with letters of recommendation. *Public Personnel Management*, 22, 81-90.
- Ackerman, S.J., & Hilsenroth, M.J. (2001). A review of therapist characteristics and techniques negatively impacting the therapeutic alliance. *Psychotherapy*, 38, 171-185.
- Ackerman, S.J., & Hilsenroth, M.J. (2003). A review of therapist characteristics and techniques positively impacting the therapeutic alliance. *Clinical Psychology Review*, 23, 1-33.
- Ahn, H., & Wampold, B.E. (2001). Where oh where are the specific ingredients? A meta-analysis of component studies in counseling and psychotherapy. *Journal of Counseling Psychology*, 48, 251-257.
- American Psychological Association. (2006a). *APA task force on the assessment of competence in professional psychology: Final report*. Washington DC: American Psychological Association.
- American Psychological Association Presidential Task Force on Evidence-Based Practice. (2006b). Evidence-based practice in psychology. *American Psychologist*, 61, 271-285.
- American Psychological Association Committee on Training in Clinical Psychology. (1947). Recommended graduate training program in clinical psychology. *American Psychologist*, 2, 539-558.
- American Psychological Association Committee on Accreditation. (2008). *Guidelines and principles for accreditation of programs in professional psychology*. Washington, DC: American Psychological Association.

- Anderson, T., Ogles, B.M., Patterson, C.L., Lambert, M.J., & Vermeersch, D.A. (2009). Therapist effects: Facilitative interpersonal skills as a predictor of therapist success. *Journal of Clinical Psychology, 65*, 755-768.
- Anderson, T., Patterson, C.L., & Weiss, A.C. (2006). Facilitative interpersonal skills performance analysis rating method. Unpublished coding manual. Department of Psychology, Ohio University, Athens, OH.
- Asay, T., & Lambert, M. (2002). Therapist relational variables. In D.J. Cain & J. Seeman (Eds.), *Humanistic psychotherapies: Handbook of research and practice* (1st ed., pp. 531-559). Washington DC: American Psychological Association.
- Assessment of Competency Benchmarks Work Group. (2007). *A developmental model for defining and measuring competence in professional psychology*. APA Board of Educational Affairs in collaboration with the Council of Chairs of Training Councils. Retrieved July 30, 2008, from http://www.apa.org/ed/graduate/comp_benchmark.pdf
- Atkins, D.C., & Christensen, A. (2001). Is professional training worth the bother? A review of the impact of psychotherapy training on client outcome. *Australian Psychologist, 36*, 122-130.
- Aveline, M. (2005). The person of the therapist. *Psychotherapy Research, 15*, 155-164.
- Bachelor, A. (1988). How clients perceive therapist empathy: A content analysis of “received” empathy. *Psychotherapy, 25*, 227-240.
- Bachelor, A. (1995). Clients’ perception of the therapeutic alliance: A qualitative analysis. *Journal of Counseling Psychology, 42*, 323-337.

- Baldwin, S.A., Wampold, B.E., & Imel, Z.E. (2007). Untangling the alliance-outcome correlation: Exploring the relative importance of therapist and patient variability in the alliance. *Journal of Consulting and Clinical Psychology, 75*, 842-852.
- Barber, J.P. (2009). Toward a working through of some core conflicts in psychotherapy research. *Psychotherapy Research, 19*, 1-12.
- Barber, J.P., Connolly, M.B., Crits-Christoph, P., Gladis, L., & Siqueland, L. (2000). Alliance predicts patients' outcome beyond in-treatment change in symptoms. *Journal of Consulting and Clinical Psychology, 68*, 1027-1032.
- Barber, J.P., Gallop, R., Crits-Christoph, P., Barrett, M., Klostermann, S., McCarthy, K.S., & Sharpless, B.A. (2008). The role of the alliance and techniques in predicting outcome of supportive-expressive dynamic therapy for cocaine dependence. *Psychoanalytic Psychology, 25*, 461-482.
- Barkham, M., & Shapiro, D.A. (1986). Counselor verbal response modes and experienced empathy. *Journal of Counseling Psychology, 33*, 3-10.
- Barrett-Lennard, G.T. (1962). Dimensions of therapist response as causal factors in therapeutic response as causal factors in therapeutic change. *Psychological Monographs, 76*, 1-33.
- Barrett-Lennard, G.T. (1981). The empathy cycle: Refinement of a nuclear concept. *Journal of Counseling Psychology, 28*, 91-100.
- Barrett-Lennard, G.T. (1993). The phases and focuses of empathy. *British Journal of Medical Psychology, 66*, 3-14.

- Baskin, T.W., Tierney, S.C., Minami, T., & Wampold, B.E. (2003). Establishing specificity in psychotherapy: A meta-analysis of structural equivalence of placebo controls. *Journal of Consulting and Clinical Psychology, 71*, 973-979.
- Bedi, R.P. (2006). Concept mapping the client's perspective on counseling alliance formation. *Journal of Counseling Psychology, 53*, 26-35.
- Behnke, S. (2005). Ethnic Rounds: The supervisor as gatekeeper. *APA Monitor, 36*, p. 1.
- Benish, S.G., Imel, Z.E., & Wampold, B.E. (2008). The relative efficacy of bona fide psychotherapies for treating post-traumatic stress disorder: A meta-analysis of direct comparisons. *Clinical Psychology Review, 28*, 746-758.
- Benjamin, L.S. (1974). Structural analysis of social behavior. *Psychological Review, 81*, 392-425.
- Bergin, A.E. (1997). Neglect of the therapist and the human dimensions of change: A commentary. *Clinical Psychology: Science and Practice, 4*, 83-89.
- Bergin, A.E., & Jasper, L.G. (1969). Correlates of empathy in psychotherapy: A replication. *Journal of Abnormal Psychology, 74*, 477-481.
- Bergin, A., & Solomon, S. (1970). Personality and performance correlates of empathic understanding in psychotherapy. In T. Tomlinson & J. Hart (Eds.), *New directions in client-centered therapy*. Boston, MA: Houghton-Mifflin.
- Berk, L.E. (2000). *Child development* (5th ed.). Needham Heights, MA: Allyn & Bacon.
- Beutler, L.E. (1979). Toward specific psychological therapies for specific conditions. *Journal of Consulting and Clinical Psychology, 47*, 882-897.
- Beutler, L.E. (1983). *Eclectic psychotherapy: A systematic approach*. New York, NY: Pergamon Press.

- Beutler, L.E. (1995). The germ theory myth and the myth of outcome homogeneity. *Psychotherapy: Theory, Research, Practice, Training*, 32, 498-494.
- Beutler, L.E. (1997). The psychotherapist as a neglected variable in psychotherapy: An illustration by reference to the role of therapist experience and training. *Clinical Psychology: Science and Practice*, 4, 44-52.
- Beutler, L.E. (2002). The dodo bird is extinct. *Clinical Psychology: Science and Practice*, 9, 30-34.
- Beutler, L.E., & Harwood, T.M. (2002). What is and can be attributed to the therapeutic relationship? *Journal of Contemporary Psychotherapy*, 32, 25-33.
- Beutler, L.E., Machado, P.P., & Neufeldt, S.A. (1994). Therapist variables. In A.E. Bergin & S.L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 229-269). New York, NY: John Wiley & Sons.
- Beutler, L.E., Malik, M., Alimohamed, S., Harwood, T.M., Talebi, H., Noble, S., & Wong, E. (2004). Therapist variables. In M.J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapy and behavior change* (5th ed., pp. 227-306). New York, NY: John Wiley & Sons, Inc.
- Biaggio, M.K., Gasparikova-Kransnec, M., & Bauer, L. (1983). Evaluation of clinical psychology graduate students: The problem of the unsuitable student. *Professional Practice of Psychology*, 4, 9-20.
- Bickman, L. (1999). Practice makes perfect and other myths about mental health services. *American Psychologist*, 54, 965-978.

- Binder, J.L., & Strupp, H.H. (1997). "Negative process": A recurrently discovered and underestimated facet of therapeutic process and outcome in the individual psychotherapy of adults. *Clinical Psychology: Science and Practice*, 4, 121-139.
- Blatt, S.J., Sanislow, C.A., Zuroff, D.C., & Pilkonis, P.A. (1996). Characteristics of effective therapists: Further analyses of data from the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology*, 64, 1276-1284.
- Blatt, S.J., Zuroff, D.C., Quinlan, D.M., & Pilkonis, P.A. (1996). Interpersonal factors in brief treatment of depression: Further analyses of the National Institute of Mental Health Treatment of Depression Collaborative Research Program. *Journal of Consulting and Clinical Psychology*, 64, 162-171.
- Bohart, A.C., Elliott, R., Greenberg, L.S., & Watson, J.C. (2002). Empathy. In J.C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 89-108). New York, NY: Oxford University Press.
- Bohart, A.C., & Greenberg, L.S. (1997a). Empathy and psychotherapy: An introductory overview. In A.C. Bohart & L.S. Greenberg (Eds.), *Empathy reconsidered: New directions in psychotherapy* (pp. 3-32). Washington DC: American Psychological Association.
- Bohart, A.C., & Greenberg, L.S. (1997b). Empathy: Where are we and where do we go from here? In A.C. Bohart & L.S. Greenberg (Eds.), *Empathy reconsidered: New directions in psychotherapy* (pp. 419-450). Washington DC: American Psychological Association.

- Borders, D.L., & Fong, M.L. (1991). Evaluations of supervisees: Brief commentary and research report. *Clinical Supervisor, 9*, 43-51.
- Bordin, E.S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research and Practice, 16*, 252-260.
- Bordin, E.S. (1983). A working alliance based model of supervision. *The Counseling Psychologist, 11*, 35-42.
- Botella, L., Corbella, S., Belles, L., Pacheco, M., Gomez, A.M., Herrero, O., Ribas, E., & Pedro, N. (2008). Predictors of therapeutic outcome and process. *Psychotherapy Research, 18*, 535-542.
- Bowlby, J. (1988). *A secure base: Parent-child attachment and healthy human development*. New York, NY: Basic Books.
- Brady, J.P., Davison, G.C., Dewald, P.A., Egan, G., Fadiman, J., Frank, J.D., Gill, M.M., Hoffman, I., Kempler, W., Lazarus, A.A., Raimy, V., Rotter, J.B., & Strupp, H.H. (1980). Some views on effective principles of psychotherapy. *Cognitive Therapy and Research, 4*, 269-306.
- Broadhurst, A. (1976). Validity of interview for selecting students for postgraduate training in clinical psychology. *Psychological Reports, 38*, 887-894.
- Brown, R.M. (2004). Self-composed: Rhetoric in psychology personal statements. *Written Communication, 21*, 242-260.
- Brown, G.S., Lambert, M.J., Jones, E.R., & Minami, T. (2005). Identifying highly effective psychotherapists in a managed care environment. *The American Journal of Managed Care, 11*, 513-520.

- Buser, T.J. (2008). Counselor training: Empirical findings and current approaches. *Counselor Education and Supervision, 48*, 86-100.
- Butler, S.F., & Strupp, H.H. (1986). Specific and nonspecific factors in psychotherapy: A problematic paradigm for psychotherapy research. *Psychotherapy, 23*, 30-39.
- Cahill, J. (1981). Towards a model refinement in compensatory education: Comparison of intervention programs and paraprofessional screening measures. *American Journal of Community Psychology, 9*, 731-749.
- Carkhuff, R.R. (1969). *Helping and human relations: A primer for lay and professional helpers (Vols. I and II)*. New York, NY: Holt, Rinehart, & Winston, Inc.
- Castonguay, L.G., & Beutler, L.E. (2006). Principles of therapeutic change: A task force on participants, relationships, and techniques factors. *Journal of Clinical Psychology, 62*, 631-638.
- Castonguay, L.G., Boswell, J.F., Constantino, M.J., Goldfried, M.R., & Hill, C.E. (2010). Training implications of harmful effects of psychological treatments. *American Psychologist, 65*, 34-49.
- Castonguay, L.G., Constantino, M.J., & Grosse Holtforth, M. (2006). The working alliance: Where are we and where should we go? *Psychotherapy: Theory, Research Practice, Training, 43*, 271-279.
- Castonguay, L.G., Goldfried, M.R., Wiser, S., Raue, P.J., & Hayes, A.M. (1996). Predicting the effect of cognitive therapy for depression: A study of unique and common factors. *Journal of Consulting and Clinical Psychology, 64*, 497-504.
- Chambless, D.L. (2002). Beware the dodo bird: The dangers of overgeneralization. *Clinical Psychology: Science and Practice, 9*, 13-16.

- Chambless, D.L., Baker, M., Baucom, D.H., Beutler, L.E., Calhoun, K.S. et al. (1998). Update on empirically validated therapies, II. *The Clinical Psychologist*, 51, 3-16.
- Chambless, D.L., & Ollendick, T.H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology*, 52, 685-716.
- Chinsky J.M., & Rappaport, J. (1971). Evaluation of a technique for the behavioral assessment of nonprofessional mental health workers. *Journal of Clinical Psychology*, 27, 400-402.
- Clarkin, J.F., & Levy, K.N. (2004). The influence of client variables on psychotherapy. In M.J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapy and behavior change* (5th ed., pp. 194-226). New York, NY: John Wiley & Sons, Inc.
- Coady, N.F. (1991a). The association between client and therapist interpersonal processes and outcomes in psychodynamic psychotherapy. *Research on Social Work Practice*, 1, 122-138.
- Coady, N.F. (1991b). The association between complex types of therapist interventions and outcomes in psychodynamic psychotherapy. *Research on Social Work Practice*, 1, 257-277.
- Constantino, M.J. (2000). Interpersonal process in psychotherapy through the lens of the structural analysis of social behavior. *Applied & Preventative Psychology*, 9, 153-172.
- Constantino, M.J., Castonguay, L.G., & Schut, A.J. (2002). The working alliance: A flagship for the “scientist-practitioner” model in psychotherapy. In G.S. Tryon (Ed.),

- Counseling based on process research: Applying what we know* (pp. 81-131). Boston, MA: Allyn & Bacon.
- Constanzo, M., & Philpott, J. (1986). Predictors of therapeutic talent in aspiring clinicians: A multivariate analysis. *Psychotherapy, 23*, 363-369.
- Coutinho, J., Ribeiro, E., Hill, C., & Safran, J. (2011). Therapists' and client's experiences of alliance ruptures: A qualitative study. *Psychotherapy Research, 21*, 525-540.
- Critelli, J.W., & Neumann, K.F. (1984). The placebo: Conceptual analysis of a construct in transition. *American Psychologist, 39*, 32-39.
- Crits-Christoph, P., Baranacki, K., Kurcias, J., Beck, A., Carroll, K., Perry, K., Luborsky, L., McClellan, A., Woody, G., Thompson, L., Gallagher, D., & Zitrin, C. (1991). Meta-analysis of therapist effects in psychotherapy outcome studies. *Psychotherapy Research, 1*, 81-91.
- Crits-Christoph, P., Connolly Gibbons, M.B., Crits-Christoph, K., Narducci, J., Schamberger, M., & Gallop, R. (2006). Can therapists be trained to improve their alliances? A preliminary study of alliance-fostering psychotherapy. *Psychotherapy Research, 16*, 268-281.
- Crits-Christoph, P., Connolly Gibbons, M.B., & Hearon, B. (2006). Does the alliance cause good outcome? Recommendations for future research on the alliance. *Psychotherapy: Theory, Research, Practice, Training, 43*, 280-285.
- Crits-Christoph, P., & Mintz, J. (1991). Implications of therapist effects for the design and analysis of comparative studies of psychotherapies. *Journal of Consulting and Clinical Psychology, 59*, 20-26.

- Cuijpers, P., van Straten, A., Andersson, G., & van Oppen, P. (2008). Psychotherapy for depression in adults: A meta-analysis of comparative outcome studies. *Journal of Consulting and Clinical Psychology, 76*, 909-922.
- Daehnert, C., & Carter, J.D. (1987). The prediction of success in a clinical psychology graduate program. *Educational and Psychological Measurement, 47*, 1113-1125.
- Daniel, S.I. (2006). Adult attachment patterns and individual psychotherapy: A review. *Clinical Psychology Review, 26*, 968-984.
- Dawes, R.M. (1975). Graduate admission variables and future success. *Science, 187*, 721-723.
- Dawes, R.M. (1994). House of cards: Psychology and psychotherapy built on myth. New York, NY: The Free Press.
- DeRubeis, R.J., Brotman, M.A., & Gibbons, C.J. (2005). A conceptual and methodological analysis of the nonspecifics argument. *Clinical Psychology: Science and Practice, 12*, 174-183.
- DeRubeis, R.J., & Feeley, M. (1990). Determinants of change in cognitive therapy for depression. *Cognitive Therapy and Research, 14*, 469-482.
- Dicken, C., Bryson, R., & Kass, N. (1977). Companionship therapy: A replication in experimental community psychology. *Journal of Consulting and Clinical Psychology, 45*, 637-646.
- Dinger, U., Strack, M., Leichsenring, F., Wilmers, F., & Schauenburg, H. (2008). Therapist effects on outcome and alliance in inpatient psychotherapy. *Journal of Clinical Psychology, 64*, 344-354.

- Dollinger, S.J. (1989). Predictive validity of the Graduate Record Examination in a clinical psychology program. *Professional Psychology: Research and Practice*, 20, 56-58.
- Dooley, D. (1975). Selecting nonprofessional counselor trainees with the group assessment of interpersonal traits (GAIT). *American Journal of Community Psychology*, 3, 371-383.
- Dooley, D. (1980). Screening of paratherapists: Empirical status and research directions. *Professional Psychology*, 11, 242-251.
- Dooley, D., Lange, A.J., & Whiteley, J.M. (1979). Sources of discrepancy in Carkhuff and GAIT measurements of empathy. *Psychotherapy: Theory, Research, and Practice*, 16, 337-344.
- Duan, C., & Hill, C.E. (1996). The current state of empathy research. *Journal of Counseling Psychology*, 43, 261-274.
- Duff, C.T., & Bedi, R.P. (2010). Counsellor behaviors that predict therapeutic alliance: From the client's perspective. *Counselling Psychology Quarterly*, 23, 91-110.
- Easton, C., Martin, W.E., & Wilson S. (2008). Counselor preparation: Emotional intelligence and implications for counseling self-efficacy: Phase II. *Counselor Education & Supervision*, 47, 218-232.
- Eaton, T., Abeles, N., & Gutreund, M. (1993). Negative indicators, therapeutic alliance, and therapy outcome. *Psychotherapy Research*, 3, 115-123.
- Eisenberg, N., & Fabes, R.A. (1990). Empathy: Conceptualization, measurement, an relation to prosocial behavior. *Motivation and Emotion*, 14, 131-149.

- Eisenberg, N., Fabes, R.A., Schaller, M., Carlo, G., & Miller, P.A. (1991). The relations of parental characteristics and practices to children's vicarious emotional responding. *Child Development, 62*, 1393-1408.
- Elkin, I., Shea, M.T., Watkins, J.T., Imber, S.D., Stosky, S.M., Collins, J.F., Glass, D.R., Pilkonis, P.A., Weber, W.R., Docherty, J.P., Fiester, S.J., & Parloff, M.B. (1989). National Institute of Mental Health treatment of Depression Collaborative Research Program: General effectiveness of treatments. *Archives of General Psychiatry, 46*, 971-982.
- Elliott, R., Bohart, A.C., Watson, J.C., & Greenberg, L.S. (2011). Empathy. *Psychotherapy, 48*, 43-49.
- Elliott, R., & James, E. (1989). Varieties of client experience in psychotherapy: An analysis of the literature. *Clinical Psychology Review, 9*, 443-468.
- Ellis, M.V., & Ladany, N. (1997). Inferences concerning supervisees and clients in clinical supervision: An integrative review. In C.E. Watkins, Jr. (Ed.), *Handbook of psychotherapy supervision* (pp. 447-507). New York, NY: John Wiley & Sons, Inc.
- Elman, N.S., Illfelder-Kaye, J., & Robiner, W.N. (2005). Professional development: Training for professionalism as a foundation for competent practice in psychology. *Professional Psychology: Research and Practice, 36*, 367-375.
- Elvins, R., & Green, J. (2008). The conceptualization and measurement of therapeutic alliance: An empirical review. *Clinical Psychology Review, 28*, 1167-1187.
- Farber, B.A., & Lane, J.S. (2002). Positive regard. In J.C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 175-194). New York, NY: Oxford University Press.

- Fauber, R.L. (2006). Graduate admissions in clinical psychology: Observations on the present and thoughts on the future. *Clinical Psychology: Science and Practice*, 13, 227-234.
- Federici, L., & Schuerger, J. (1974). Prediction of success in an applied M.A. psychology program. *Educational and Psychological Measurement*, 34, 945-952.
- Feeley, M., DeRubeis, R.J., & Gefland, L.A. (1999). The temporal relation of adherence and alliance to symptom change in cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, 67, 578-582.
- Feshbach, N.D. (1997). Empathy: The formative years – implications for clinical practice. In A.C. Bohart & L.S. Greenberg (Eds.), *Empathy reconsidered: New directions in psychotherapy* (pp. 33-62). Washington DC: American Psychological Association.
- Fitzpatrick, M.R., Stalikas, A., & Iwakabe, S. (2001). Examining counselor interventions and client progress in the context of the therapeutic alliance. *Psychotherapy*, 38, 160-170.
- Forrest, L., Elman, N., Gizara, S., & Vacha-Haase, T. (1999). Trainee impairment: A review of identification, remediation, dismissal, and legal issues. *The Counseling Psychologist*, 27, 627-686.
- Frank, J.D., & Frank, J.B. (1991). *Persuasion and healing: A comparative study of psychotherapy* (3rd ed.). Baltimore, MD: John Hopkins University Press.
- Gaffan, E.A., Tsauousis, I., & Kemp-Wheeler, S.M. (1995). Researcher allegiance and meta-analysis: The case of cognitive therapy for depression. *Journal of Consulting and Clinical Psychology*, 63, 966-980.

- Garfield, S.L. (1973). Basic ingredients or common factors in psychotherapy? *Journal of Consulting and Clinical Psychology*, 41, 9-12.
- Garfield, S.L. (1995). *Psychotherapy: An eclectic-integrative approach*. New York, NY: Wiley & Sons.
- Gaston, L. (1990). The concept of the alliance and its role in psychotherapy: Theoretical and empirical considerations. *Psychotherapy*, 27, 143-153.
- Gaston, L., Piper, W., Debbane, E., Bienvenu, J., & Garant, J. (1994). Alliance and technique for predicting outcome in short-and long-term analytic psychotherapy. *Psychotherapy Research*, 4, 121-135.
- Gaston, L., Thompson, L., Gallagher, D., Cournoyer, L., & Gagnon, R. (1998). Alliance, technique, and their interactions in predicting outcome of behavioral, cognitive, and brief dynamic therapy. *Psychotherapy Research*, 8, 190-209.
- GlenMaye, L., & Oaks, M. (2000). Assessing suitability of MSW applicants through objective scoring of personal statements. *Journal of Social Work Education*, 38, 67-82.
- Goldfried, M.R. (1980). Toward the delineation of therapeutic change principles. *American Psychologist*, 35, 991-999.
- Goldman, R.N., Greenberg, L.S., & Angus, L. (2006). The effects of adding emotion-focused interventions to the client-centered relationship conditions in the treatment of depression. *Psychotherapy Research*, 16, 537-549.
- Gonsalvez, C.J., & Freestone, J. (2007). Field supervisors' assessments of trainee performance: Are they reliable and valid? *Australian Psychologist*, 42, 23-32.

- Goodheart, C.D., Kazdin, A.E., & Sternberg, R.J. (Eds.). (2006). *Evidence-based psychotherapy: Where practice and research meet*. Washington DC: American Psychological Association.
- Goodman, G. (1972). *Companionship therapy*. San Francisco, CA: Jossey-Bass, Inc.
- Gormally, J., & Hill, C.E. (1974). Guidelines for research on Carkhuff's training model. *Journal of Counseling Psychology, 21*, 539-547.
- Gray, L.A., Ladany, N., Walker, J.A., & Ancis, J.R. (2001). Psychotherapy trainees' experience of counterproductive events in supervision. *Journal of Counseling Psychology, 48*, 371-383.
- Grencavage, L.M., & Norcross, J.C. (1990). Where are the commonalities among the therapeutic common factors? *Professional Psychology: Research and Practice, 21*, 372-378.
- Grissom, R.J. (1996). The magical number .7 +/- .2: Meta-meta-analysis of the probability of superior outcome in comparisons involving therapy, placebo, and control. *Journal of Consulting and Clinical Psychology, 64*, 973-982.
- Grote, C.L., Robiner, W.N., & Haught, A. (2001). Disclosure of negative information in letters of recommendation: Writers' intentions and readers' experiences. *Professional Psychology: Research and Practice, 32*, 655-661.
- Gurman, A.S. (1977). The patient's perception of the therapeutic relationship. In A.S. Gurman, & A.M. Razin (Eds.), *Effective psychotherapy: A handbook of research* (pp. 503-543). Elmsford, NY: Pergamon Press.
- Gurman, A.S., & Razin, A.M. (Eds.). (1977). *Effective psychotherapy: A handbook of research*. Elmsford, NY: Pergamon Press.

- Hackman, J.R., Wiggins, N., & Bass, A.R. (1970). Prediction of long-term success in doctoral work in psychology. *Educational and Psychological Measurement, 30*, 365-374.
- Harmon, S.C., Lambert, M.J., Smart, D.M., Hawkins, E., Nielsen, S.L., Slade, K., & Lutz, W. (2007). Enhancing outcome for potential treatment failures: Therapist-client feedback and clinical support tools. *Psychotherapy Research, 17*, 379-392.
- Hatcher, R.L., & Barends, R.L. (2006). How a return to theory could help alliance research. *Psychotherapy: Theory, Research, Practice, Training, 43*, 292-299.
- Hatcher, R.L., & Lassiter, K.D. (2007). Initial training in professional psychology: The practicum competencies outline. *Training and Education in Professional Psychology, 1*, 49-63.
- Henry, W.P. (1996). Structural analysis of social behavior as a common metric for programmatic psychopathology and psychotherapy research. *Journal of Consulting and Clinical Psychology, 64*, 1263-1275.
- Henry, W.P., Schacht, T.E., & Strupp, H.H. (1986). Structural analysis of social behavior: Application to a study of interpersonal process in differential psychotherapeutic outcome. *Journal of Consulting and Clinical Psychology, 54*, 27-31.
- Henry, W.P., Schacht, T.E., & Strupp, H.H. (1990). Patient and therapist introject, interpersonal process, and differential psychotherapy outcome. *Journal of Consulting and Clinical Psychology, 58*, 768-774.

- Henry, W.P., Schacht, T.E., Strupp, H.H., Butler, S.F., & Binder, J.L. (1993). Effects of training in time-limited dynamic psychotherapy: Mediators of therapists response to training. *Journal of Consulting and Clinical Psychology, 61*, 441-447.
- Henry, W.P., & Strupp, H.H. (1994). The therapeutic alliance as an interpersonal process. In A.O. Horvath & L.S. Greenberg (Eds.), *The working alliance: Theory, research, and practice* (pp. 51-84). New York, NY: John Wiley & Sons.
- Henry, W.P., Strupp, H.H., Butler, S.F., Schacht, T.E., & Binder, J.L. (1993). Effects of training in time-limited dynamic psychotherapy: Changes in therapist behavior. *Journal of Consulting and Clinical Psychology, 61*, 434-440.
- Hill, C.E. (Ed.). (2001). *Helping skills: The empirical foundation*. Washington DC: American Psychological Association.
- Hill, C.E. (2004). *Helping skills: Facilitating exploration, insight, and action* (2nd ed.). Washington DC: American Psychological Association.
- Hill, C.E. (2005). Therapist techniques, client involvement, and the therapeutic relationship: Inextricably intertwined in the therapy process. *Psychotherapy: Theory, Research, Practice, Training, 42*, 431-442.
- Hill, C.E., Kellems, I.S., Kolchakian, M.R., Wonnell, T.L., Davis, T.L., & Nakayama, E.Y. (2003). The therapist experience of being the target of hostile versus suspected-unasserted client anger: Factors associated with resolution. *Psychotherapy Research, 13*, 475-491.
- Hill, C.E., & Knox, S. (2009). Processing the therapeutic relationship. *Psychotherapy Research, 19*, 13-29.

- Hill, C.E., & Lambert, M.J. (2004). Methodological issues in studying psychotherapy processes and outcomes. In M.J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapy and behavior change* (5th ed., pp. 84-136). New York, NY: John Wiley & Sons, Inc.
- Hill, C.E., Roffman, M., Stahl, J., Friedman, S., Hummel, A., & Wallace, C. (2008). Helping skills training for undergraduates: Outcomes and prediction of outcomes. *Journal of Counseling Psychology*, 55, 359-370.
- Hill, C.E., Thompson, B.J., Cogar, M.C., & Denman, D.W. (1993). Beneath the surface of long-term therapy: Therapist and client report of their own and each other's covert processes. *Journal of Consulting Psychology*, 40, 278-287.
- Hill, C., Thompson, B., & Corbett, M. (1992). The impact of therapist ability to perceive displayed and hidden client reactions on immediate outcome in first sessions of brief therapy. *Psychotherapy Research*, 2, 143-155.
- Hilliard, R.B., Henry, W.P., & Strupp, H.H. (2000). An interpersonal model of psychotherapy: Linking patient and therapist developmental history, therapeutic process, and types of outcome. *Journal of Consulting and Clinical Psychology*, 68, 125-133.
- Hilsenroth, M.J., Ackerman, S.J., Clemence, A.J., Strassle, C.G., & Handler, L. (2002). Effects of structured clinician training on patient and therapist perspectives of alliance early in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, 39, 309-323.
- Hirschberg, N., & Itkin, S. (1978). Graduate student success in psychology. *American Psychologist*, 33, 1083-1093.

- Hoffart, A., Borge, F., Sexton, H., & Clark, D.M. (2009). The role of common factors in residential cognitive and interpersonal therapy for social phobia: A process-outcome study. *Psychotherapy Research, 19*, 54-67.
- Hoffart, A., Sexton, H., Nordahl, H.M., & Stiles, T.C. (2005). Connection between patient and therapist and therapist's competence in schema-focused therapy of personality problems. *Psychotherapy Research, 15*, 409-419.
- Horvath, A.O. (2006). The alliance in context: Accomplishments, challenges, and future directions. *Psychotherapy: Theory, Research, Practice, Training, 43*, 258-263.
- Horvath, A.O., & Bedi, R.P. (2002). The alliance. In J.C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 37-70). New York, NY: Oxford University Press.
- Horvath, A.O., & Greenberg, L.S. (1986). Development of the working alliance inventory. In L.S. Greenberg, & W.M. Pinsof (Eds.), *The psychotherapeutic process: A research handbook* (pp. 529-556). New York, NY: Guilford Press.
- Horvath, A.O., & Luborsky, L. (1993). The role of the therapeutic alliance in psychotherapy. *Journal of Consulting and Clinical Psychology, 61*, 561-573.
- Horvath, A.O., & Symonds, B.D. (1991). Relation between working alliance and outcome in psychotherapy: A meta-analysis. *Journal of Counseling Psychology, 38*, 139-149.
- Hosford, R.E., Johnson, M.E., & Atkinson, D.R. (1984). Academic criteria, experiential background, and personal interviews as predictors of success in a counselor education program. *Counselor Education and Supervision, 23*, 268-275.

- Hrobjartsson, A., & Gotzsche, P.C. (2007). Powerful spin in the conclusion of Wampold et al.'s re-analysis of placebo versus no-treatment trials despite similar results as in original review. *Journal of Clinical Psychology, 63*, 373-377.
- Huitema, B.E., & Stein, C.R. (1993). Validity of the GRE without restriction of range. *Psychological Reports, 72*, 123-127.
- Hunsley, J., & Westmacott, R. (2007). Interpreting the magnitude of the placebo effect: Mountain or molehill? *Journal of Clinical Psychology, 63*, 391-399.
- Huppert, J.D., Bufka, L.F., Barlow, D.H., Gorman, J.M., & Shear, M.K. (2001). Therapists, therapist variables, and cognitive-behavioral therapy outcome in a multicenter trial for panic disorder. *Journal of Consulting and Clinical Psychology, 69*, 747-755.
- Hurst, M.W., & Shatkin, S.D. (1974). Relationship between standardized admissions variables and certain interpersonal skills. *Counselor Education and Supervision, 14*, 22-33.
- Imber, S.D., Pilkonis, P.A., Sotsky, S.M., Elkin, I., Watkins, J.T., Collins, J.F., Shea, M.T., Leber, W.R., & Glass, D.R. (1990). Mode-specific effects among three treatments for depression. *Journal of Consulting and Clinical Psychology, 58*, 352-359.
- Ingram, R.E. (1983). The GRE in the graduate admissions process: Is how it is used justified by the evidence of its validity? *Professional Psychology: Research and Practice, 14*, 711-714.
- Ingram, R.E., & Zurawski, R. (1981). Choosing clinical psychologists: An examination of the utilization of admissions criteria. *Professional Psychology, 12*, 684-689.

- Jackson, E. (1986). Affective sensitivity and peer selection of counseling potential. *Counselor Education & Supervision, 25*, 230-236.
- Jacobson, N.S., & Truax, P. (1991). Clinical significance: A statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology, 59*, 12-19.
- Jennings, L., & Skovholt, T.M. (1999). The cognitive, emotional, and relational characteristics of master therapists. *Journal of Counseling Psychology, 46*, 3-11.
- Johnson, W.B., & Campbell, C.D. (2002). Character and fitness requirements for professional psychologists: Are there any? *Professional Psychology: Research and Practice, 33*, 46-53.
- Johnson, W.B., & Campbell, C.D. (2004). Character and fitness requirements for professional psychologists: Training directors' perspectives. *Professional Psychology: Research and Practice, 35*, 405-411.
- Johnson, D.W., & Matross, R.P. (1977). Interpersonal influence in psychotherapy: A social psychological view. In A.S. Gurman & A.M. Razin (Eds.), *Effective psychotherapy: A handbook of research* (pp. 395-432). Elmsford, NY: Pergamon Press.
- Karasu, T.B. (1986). The specificity versus nonspecificity dilemma: Toward identifying therapeutic change agents. *American Journal of Psychiatry, 143*, 687-695.
- Kaslow, N.J., Pate, W.E., & Thorn, B. (2005). Academic and internship directors' perspectives on practicum experiences: Implications for training. *Professional Psychology: Research and Practice, 36*, 307-317.

- Kaslow, N.J., Rubin, N.J., Forrest, L., Elman, N.S., VanHorne, B.A., Jacobs, S.C. et al. (2007). Recognizing, assessing, and intervening with problems of professional competence. *Professional Psychology: Research and Practice*, 38, 479-492.
- Katz, R.L. (1963). *Empathy: Its nature and uses*. London: Free Press of Glencoe.
- Kazdin, A.E. (2005). Treatment outcomes, common factors, and continued neglect of mechanisms of change. *Clinical Psychology: Science and Practice*, 12, 184-188.
- Kazdin, A.E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology*, 3, 1-27.
- Kazdin, A.E. (2008). Evidence-based treatment and practice: New opportunities to bridge clinical research and practice, enhance the knowledge base, and improve patient care. *American Psychologist*, 63, 146-159.
- Kazdin, A.E. (2009). Bridging science and practice to improve patient care. *American Psychologist*, 64, 276-279.
- Kelly, E.L., & Fiske, D.W. (1951). *The prediction of performance in clinical psychology*. Ann Arbor, MI: The University of Michigan Press.
- Kiesler, D.J. (1996). *Contemporary interpersonal theory and research: Personality, psychopathology, and psychotherapy*. New York, NY: John Wiley & Sons.
- Kiesler, D.J., & Watkins, L.M. (1989). Interpersonal complementarity and the therapeutic alliance: A study of relationship in psychotherapy. *Psychotherapy*, 26, 183-194.
- Kim, D., Wampold, B.E., & Bolt, D.M. (2006). Therapist effects in psychotherapy: A random-effects modeling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. *Psychotherapy Research*, 16, 161-172.

- King, D.W., Beehr, T.A., & King, L.A. (1986). Doctoral student selection in one professional psychology program. *Journal of Clinical Psychology, 42*, 399-407.
- Kirnan, J.P., & Geisinger, K.F. (1981). The prediction of graduate school success in psychology. *Educational and Psychological Measurement, 41*, 815-820.
- Kivlighan, D.M., & Schmitz, P.J. (1992). Counselor technical activity in cases with improving working alliances and continuing-poor working alliances. *Journal of Counseling Psychology, 39*, 32-28.
- Kivlighan, D.M., & Shaughnessy, P. (2000). Patterns of working alliance development: A typology of client's working alliance ratings. *Journal of Counseling Psychology, 47*, 362-371.
- Klee, M.R., Abeles, N., & Muller, R.T. (1990). Therapeutic alliance: Early indicators, course, and outcome. *Psychotherapy, 27*, 166-173.
- Klein, D.N., Schwartz, J.E., Santiago, N.J., Vivian, D., Vocisano, C., Castonguay, L.G., Arnow, B., Blalock, J.A., Manber, R., Markowitz, J.C., Riso, L.P., Rothbaum, B., McCullough, J.P., Thase, M.E., Borian, F.E., Miller, I.W., & Keller, M.B. (2003). Therapeutic alliance in depression treatment: Controlling for prior change and patient characteristics. *Journal of Consulting and Clinical Psychology, 71*, 997-1006.
- Ko, S.F., & Rodolfa, E. (2005). Psychology training directors' views of number of practicum hours necessary prior to internship application. *Professional Psychology: Research and Practice, 36*, 318-322.
- Korman, M. (1974). National conference on levels and patterns of professional training in psychology: The major themes. *American Psychologist, 29*, 441-449.

- Krause, M.S., & Lutz, W. (2009). Process transforms inputs to determine outcomes: Therapists are responsible for managing process. *Clinical Psychology: Science and Practice*, 16, 73-81.
- Krause, M.S., Lutz, W., & Saunders, S.M. (2007). Empirically certified treatments or therapists: The issue of separability. *Psychotherapy: Theory, Research, Practice, Training*, 44, 347-353.
- Kuncel, N.R., & Hezlett, S.A. (2007). Standardized tests predict graduate students' success. *Science*, 315, 1080-1081.
- Kuncel, N.R., Hezlett, S.A., & Ones, D.S. (2001). A comprehensive meta-analysis of the predictive validity of the Graduate Record Examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, 127, 162-181.
- Kurtz, R.R., & Grummon, D.L. (1972). Different approaches to the measurement of therapist empathy and their relationship to therapy outcomes. *Journal of Consulting and Clinical Psychology*, 39, 106-115.
- Ladany, N. (2007). Does psychotherapy training matter? Maybe not. *Psychotherapy: Theory, Research, Practice, Training*, 44, 392-396.
- Ladany, N., Friedlander, M.L., & Nelson, M.L. (2005). *Critical events in psychotherapy supervision: An interpersonal approach*. Washington DC: American Psychological Association.
- Lafferty, P., Beutler, L.E., & Crago, M. (1989). Differences between more and less effective psychotherapists: A study of select therapist variables. *Journal of Consulting and Clinical Psychology*, 57, 76-80.

- Lambert, M.J. (1989). The individual therapist's contribution to psychotherapy process and outcome. *Clinical Psychology Review*, 9, 469-485.
- Lambert, M.J. (2005). Early response in psychotherapy: Further evidence for the importance of common factors rather than "placebo effects." *Journal of Clinical Psychology*, 61, 855-869.
- Lambert, M.J. (2007). Presidential address: What we have learned from a decade of research aimed at improving psychotherapy outcome in routine care. *Psychotherapy Research*, 17, 1-14.
- Lambert, M.J., & Archer, A. (2006). Research findings on the effects of psychotherapy and their implications for practice. In C.D. Goodheart, A.E. Kazdin, & R.J. Sternberg (Eds.), *Evidence-based psychotherapy: Where practice and research meet* (pp. 111-130). Washington DC: American Psychological Association.
- Lambert, M.J., & Baldwin, S.A. (2009). Some observations on studying therapists instead of treatment packages. *Clinical Psychology: Science and Practice*, 16, 82-85.
- Lambert, M.J., & Barley, D.E. (2002). Research summary on the therapeutic relationship and psychotherapy outcome. In J.C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients* (pp. 17-36). New York, NY: Oxford University Press.
- Lambert, M.J., Burlingame, G.M., Umphress, V., Hansen, N.B., Vermeersch, D.A., Clouse, G.C. et al. (1996). The reliability and validity of the Outcome Questionnaire. *Clinical Psychology and Psychotherapy*, 3, 249-258.

- Lambert, M.J., & DeJulio, S.S. (1977). Outcome research in Carkhuff's human resource development training programs: Where is the donut? *The Counseling Psychologist*, 6, 79-86
- Lambert, M.H., DeJulio, S.S., & Stein, D.M. (1978). Therapist interpersonal skills: Process, outcome, methodological considerations, and recommendations for future research. *Psychological Bulletin*, 85, 467-489.
- Lambert, M.J., Garfield, S.L., & Bergin, A.E. (2004). Overview, trends, and future issues. In M.J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapy and behavior change* (5th ed., pp. 805-822). New York, NY: John Wiley & Sons, Inc.
- Lambert, M.J., Morton, J.J., Hatfield, D., Harmon, C., Hamilton, S., Reid, R.C. et al. (2004). *Administration and scoring manual for the Outcome Questionnaire-45*. Orem, UT: American Professional Credentialing Services.
- Lambert, M.J., & Ogles, B.M. (2004). The efficacy and effectiveness of psychotherapy. In M.J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapy and behavior change* (5th ed., pp. 139-193). New York, NY: John Wiley & Sons, Inc.
- Lambert, M.J., & Okiishi, J.C. (1997). The effects of the individual psychotherapist and implications for future research. *Clinical Psychology: Science and Practice*, 4, 66-75.
- Lampropoulos, G.K. (2001). Common processes of change in psychotherapy and seven other social interactions. *British Journal of Guidance and Counselling*, 29, 21-33.
- Leary, T. (1957). *Interpersonal diagnosis of personality: A functional theory and methodology for personality evaluation*. New York, NY: The Ronald Press Company.

- Lewis, B.L., Hatcher, R.L., & Pate, W.E. (2005). The practicum experience: A survey of practicum site coordinators. *Professional Psychology: Research and Practice*, 36, 291-298.
- Lichtenberg, J.W., Portnoy, S.M., Bebeau, M.J., Leigh, I.W., Nelson, P.D., Rubin, N.J., Smith, I.L., & Kaslow, N.J. (2007). Challenges to the assessment of competence and competencies. *Professional Psychology: Research and Practice*, 38, 474-478.
- Linn, R.L. (1990). Admissions testing: Recommended uses, validity, differential prediction and coaching. *Applied Measurement in Education*, 3, 297-318.
- Lipsey, M.W., & Wilson, D.B. (1993). The efficacy of psychological, educational, and behavioral treatment: Confirmation from meta-analysis. *American Psychologist*, 48, 1181-1209.
- Littlepage, G.E., Bragg, D.M., & Rust, J.O. (1978). Relations between admission criteria, academic performance, and professional performance. *Teaching of Psychology*, 5, 16-20.
- Loo, C. (1979). Measures for predicting therapeutic skill and the effects of training. *Psychotherapy: Theory, Research and Practice*, 16, 460-466.
- Luborsky, L. (1994). Therapeutic alliances as predictors of psychotherapy outcomes: Factors explaining the predictive success. In A.O. Horvath & L.S. Greenberg (Eds.), *The working alliance: Theory, research, and practice* (pp. 38-50). New York, NY: John Wiley & Sons.
- Luborsky, L., Chandler, M., Auerbach, A.H., Cohen, J., & Bachrach, H.M. (1971). Factors influencing the outcome of psychotherapy: A review of quantitative research. *Psychological Bulletin*, 75, 145-185.

- Luborsky, L., Crits-Christoph, P., McLellan, A.T., Woody, G., Piper, W., Liberman, B., Imber, S., & Pilkonis, P. (1986). Do therapists vary much in their success? Findings from four outcome studies. *American Journal of Orthopsychiatry*, 56, 501-512.
- Luborsky, L., Diguer, L., Seligman, D.A., Rosenthal, R., Krause, E.D., Johnson, S., Halperin, G., Bishop, M., Berman, J.S., & Schweizer, E. (1999). The researcher's own therapy allegiances: A "wild card" in comparisons of treatment efficacy. *Clinical Psychology: Science and Practice*, 6, 95-106.
- Luborsky, L., McLellan, A.T., Diguer, L., Woody, G., & Seligman, D.A. (1997). The psychotherapist matters: Comparison of outcomes across twenty-two therapists and seven patient samples. *Clinical Psychology: Science and Practice*, 4, 53-65.
- Luborsky, L., McLellan, A.T., Woody, G.E., O'Brien, C.P., & Auerbach, A. (1985). Therapist success and its determinants. *Archives of General Psychiatry*, 42, 602-611.
- Luborsky, L., Rosenthal, R., Diguer, L., Andrusyna, T.P., Berman, J.S., Levitt, J.T., Seligman, D.A., & Krause, E.D. (2002). The dodo bird verdict is alive and well – mostly. *Clinical Psychology: Science and Practice*, 9, 2-12.
- Luborsky, L., Singer, J., & Luborsky, L. (1975). Comparative studies of psychotherapy. *Archives of General Psychiatry*, 32, 995-1008.
- Lutz, W., Leon, S.C., Martinovich, Z., Lyons, J.S., & Stiles, W.B. (2007). Therapist effects in outpatient psychotherapy: A three-level growth curve approach. *Journal of Counseling Psychology*, 54, 32-39.
- Mallinckrodt, B., & Nelson, M.L. (1991). Counselor training level and the formation of the psychotherapeutic working alliance. *Journal of Counseling Psychology*, 38, 133-138.

- Martin, W.E., Easton, C., Wilson, S., Takemoto, M., & Sullivan, S. (2004). Counselor preparation: Salience of emotional intelligence as a core characteristic of being a counselor. *Counselor Education & Supervision, 44*, 17-30.
- Martin, D.J., Garske, J.P., & Davis, M.K. (2000). Relation of the therapeutic alliance with outcome and other variables: A meta-analytic review. *Journal of Consulting and Clinical Psychology, 68*, 438-450.
- Mayne, T.J., Norcross, J.C., & Sayette, M.A. (1994). Admission requirements, acceptance rates, and financial assistance in clinical psychology programs: Diversity of across the practice-research continuum. *American Psychologist, 49*, 806-811.
- McHolland, J., Peterson, D.R., & Brown, S.W. (1987). Assessing skills in professional psychology: A triadic strategy for developing methods. In E.F. Bourg et al. (Eds.), *Standards and evaluation in the education and training of professional psychologists* (pp. 105-128). Norman, OK: Transcript Press.
- McLellan, A.T., Woody, G.E., Luborsky, L., & Goehl, L. (1988). Is the counselor an “active ingredient” in substance abuse rehabilitation? An examination of treatment success among four counselors. *The Journal of Nervous and Mental Disease, 176*, 423-430.
- Meehl, P.E. (1957). When shall we use our heads instead of the formula? *Journal of Counseling Psychology, 4*, 268-273.
- Messer, S.B. (2004). Evidence-based practice: Beyond empirically supported treatments. *Professional Psychology: Research and Practice, 35*, 580-588.
- Miller, R.K., & Van Rycroek, G.J. (1988). Internship letters of recommendation: Where are the other 90%? *Professional Psychology: Research and Practice, 19*, 115-117.

- Missirlian, T.M., Toukmanian, S.G., Warwar, S.H., & Greenberg, L.S. (2005). Emotional arousal, client perceptual processing, and the working alliance in experiential psychotherapy for depression. *Journal of Consulting and Clinical Psychology, 73*, 861-871.
- Mitchell, K.M., Bozarth, J.D., & Krauft, C.C. (1977). A reappraisal of the therapeutic effectiveness of accurate empathy, nonpossessive warmth and genuineness. In A.S. Gurman & A.M. Razin (Eds.), *Effective psychotherapy: A handbook of research* (pp. 482-502). Elmsford, NY: Pergamon Press.
- Mohr, D.C. (1995). Negative outcome in psychotherapy: A critical review. *Clinical Psychology: Science and Practice, 2*, 1-27.
- Muran, J.C., Safran, J.D., Samstag, L.W., & Winston, A. (2005). Evaluating an alliance-focused treatment for personality disorders. *Psychotherapy: Theory, Research, Practice, Training, 42*, 532-545.
- Najavits, L.M., & Strupp, H.H. (1994). Differences in the effectiveness of psychodynamic therapists: A process-outcome study. *Psychotherapy, 31*, 114-123.
- Nathan, P.E., & Gorman, J.M. (Eds.). (2002). *A guide to treatments that work* (2nd ed.). New York, NY: Oxford University Press.
- National Council of Schools and Programs in Professional Psychology. (2007). Competency developmental achievement levels (DALs). Retrieved May 17, 2009, from <http://www.ncspp.info>
- Nevid, J.S., & Gildea, T.J. (1984). The admissions process in clinical training: The role of the personal interview. *Professional Psychology: Research and Practice, 15*, 18-25.

- Nicklin, J.M., & Roch, S.G. (2009). Letters of recommendation: Controversy and consensus from expert perspectives. *International Journal of Selection and Assessment*, 17, 76-91.
- Norcross, J.C. (2001). Purposes, processes, and products of the Task Force on Empirically Supported Therapy Relationships. *Psychotherapy*, 38, 345-356.
- Norcross, J.C. (Ed.). (2002). *Psychotherapy relationships that work: Therapist contributions and responsiveness to patients*. New York, NY: Oxford University Press.
- Norcross, J.C., Kohout, J.L., & Wickershi, M. (2005). Graduate study in psychology: 1971 to 2004. *American Psychologist*, 60, 959-975.
- Obegi, J.H. (2008). The development of the client-therapist bond through the lens of attachment theory. *Psychotherapy Theory, Research, Practice, Training*, 45, 431-446.
- O'Donovan, A., & Dyck, M. (2001). Effective training in clinical and counseling psychology: Not as simple as it sounds. *Australian Psychologist*, 36, 92-98.
- Okiishi, J.C., Lambert, M.J., Eggett, D., Nielsen, L., & Dayton, D.D. (2006). An analysis of therapist treatment effects: Toward providing feedback to individual therapists on their clients' psychotherapy outcome. *Journal of Clinical Psychology*, 62, 1157-1172.
- Okiishi, J., Lambert, M.J., Nielsen, S.L., & Ogles, B.M. (2003). Waiting for supershrink: An empirical analysis of therapist effects. *Clinical Psychology and Psychotherapy*, 10, 361-373.
- Omizo, M.M., & Michael, W.B. (1979). The prediction of performance in a counselor education master's degree program. *Educational and Psychological Measurement*, 39, 433-437.

- Orlinsky, D.E., Grawe, K., & Parks, B.K. (1994). Process and outcome in psychotherapy – Noch einmal. In A.E. Bergin & S.L. Garfield (Eds.), *Handbook of psychotherapy and behavior change* (4th ed., pp. 270-378). New York, NY: John Wiley & Sons.
- Orlinsky, D., & Howard, K.E. (1977). The therapist's experience of psychotherapy. In A.S. Gurman & A.M. Razin (Eds.), *Effective psychotherapy: A handbook of research* (pp. 566-589). Elmsford, NY: Pergamon Press.
- Orlinsky, D., & Howard, K. (1980). Gender and psychotherapeutic outcome. In A.M. Bodsky, & R.T., Hare-Mustin (Eds.), *Women and psychotherapy: An assessment of research and practice* (pp. 3-34). New York, NY: Guildford Press.
- Orlinsky, D.E., & Howard, K.I. (1986). Process and outcome in psychotherapy. In S.L. Garfield & A.E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (3rd ed.). New York, NY: John Wiley & Sons.
- Orlinsky, D.E., Ronnestad, M.H., & Willutzki, U. (2004). Fifty years of psychotherapy process-outcome research: Continuity and change. In M.J. Lambert (Ed.), *Bergin and Garfield's Handbook of psychotherapy and behavior change* (5th ed., pp. 307-390). New York, NY: John Wiley & Sons, Inc.
- Osipow, S.H., & Walsh, W.B. (1973). Social intelligence and the selection of counselors. *Journal of Counseling Psychology*, 20, 366-369.
- Parloff, M.R. (1986). Placebo controls in psychotherapy research: A sine qua non or a placebo for research problems? *Journal of Consulting and Clinical Psychology*, 54, 79-87.

- Patterson, C.L., Uhlin, B., & Anderson, T. (2008). Clients' pretreatment counseling expectation as predictors of the working alliance. *Journal of Counseling Psychology*, 55, 528-534.
- Peterson, D.R. (2003). Unintended consequences: Ventures and misadventures in the education of professional psychologists. *American Psychologist*, 58, 791-800.
- Piercy, F.P., Dickey, M., Case, B., Sprenkle, D., Beer, J., Nelson, T., & McCollum, E. (1995). Admissions criteria as predictors of performance in a family therapy doctoral program. *The American Journal of Family Therapy*, 23, 251-259.
- Piper, W.E., Azim, H.F., Joyce, A.S., McCallum, M., Nixon, G.W., & Segal, P.S. (1991). Quality of objective relationships versus interpersonal functioning as predictors of therapeutic alliance and psychotherapy outcome. *The Journal of Nervous and Mental Disease*, 179, 432-438.
- Piper, W.E., Boroto, D.R., Joyce, A.S., McCallum, M. & Azim, H.F. (1995). Pattern of alliance and outcome in short-term individual psychotherapy. *Psychotherapy*, 32, 639-647.
- Piper, W.E., Ogrodniczuk, J.S., Joyce, A.S., McCallum, M., Rosie, J.S., O'Kelly, J., & Steinberg, P.I. (1999). Prediction of dropping out in time-limited interpretive individual psychotherapy. *Psychotherapy*, 36, 114-122.
- Powers, D.E., & Fowles, M.E. (1997). The personal statement as an indicator of writing skill: A cautionary note. *Educational Assessment*, 4, 75-87.
- Principe, J.M., Marci, C.D., Glick, D.M., & Ablon, J.S. (2006). The relationship among patient contemplation, early alliance, and continuation in psychotherapy. *Psychotherapy: Theory, Research, Practice, Training*, 43, 238-243.

- Project Match Research Group. (1998). Therapist effects in three treatments for alcohol problems. *Psychotherapy Research*, 8, 455-474.
- Ramos-Sanchez, L., Esnil, E., Goodwin, A., Riggs, S., Touster, L.O., Wright, L.K., Ratanasiripoing, P., & Rodolfa, E. (2002). Negative supervisory events: Effects on supervision satisfaction and supervisory alliance. *Professional Psychology: Research and Practice*, 33, 197-202.
- Rappaport, J., Chinsky, J., & Cowen, E. (1971). *Innovations in helping chronic patients*. New York, NY: Academic Press.
- Rector, N.A., Zuroff, D.C., & Seagal, Z.V. (1999). Cognitive change and the therapeutic alliance: The role of technical and nontechnical factors in cognitive therapy. *Psychotherapy*, 36, 320-328.
- Rikard, H.C., & Clements, C.B. (1986). An evaluation of interviewed versus noninterviewed clinical psychology students. *Professional Psychology: Research and Practice*, 17, 78-79.
- 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, 355-361.
- Rodolfa, E.R., Owen, J.J., & Clark, S. (2007). Practicum training hours: Fact and fantasy. *Training and Education in Professional Psychology*, 1, 64-73.
- Rogers, C.R. (1957/2007). The necessary and sufficient conditions of therapeutic personality change. *Psychotherapy: Theory, Research, Practice, Training*, 44, 240-248.

- Rønnestad, M.H., & Ladany, N. (2006). The impact of psychotherapy training: Introduction to the special section. *Psychotherapy Research, 16*, 261-267.
- Rosenzweig, S. (1936/2002). Some implicit common factors in diverse methods of psychotherapy. *Journal of Psychotherapy Integration, 12*, 5-9.
- Rounsaville, B.J., Chevron, E.S., Prusoff, B.A., Elkin, I., Imber, S., Sotsky, S., & Watkins, J. (1987). The relation between specific and general dimensions of the psychotherapy process in interpersonal psychotherapy of depression. *Journal of Consulting and Clinical Psychology, 55*, 379-384.
- Rubin, N., Bebeau, M., Leigh, I., Lichtenberg, J., Nelson, P., Portnoy, S. et al. (2007). The competency movement within psychology: An historical perspective. *Professional Psychology: Research and Practice, 38*, 452-462.
- Safran, J.D., Crocker, P., McMain, S., & Murray, P. (1990). Therapeutic alliance rupture as a therapy event for empirical investigation. *Psychotherapy, 27*, 154-164.
- Safran, J.D., & Muran, J.C. (2000). *Negotiating the therapeutic alliance: A relational treatment guide*. New York, NY: Guilford Press.
- Safran, J.D., & Muran, J.C. (2006). Has the concept of the therapeutic alliance outlived its usefulness? *Psychotherapy: Theory, Research, Practice, Training, 43*, 286-291.
- Safran, J.D., Muran, J.C., & Samstag, L.W. (1994). Resolving therapeutic alliance ruptures: A task analytic investigation. In A.O. Horvath & L.S. Greenberg (Eds.), *The working alliance: Theory, research, and practice* (pp. 225-258). New York, NY: John Wiley & Sons.
- Safran, J.D., Muran, J.C., Samstag, L.W., & Stevens, C. (2002). Repairing alliance ruptures. In J.C. Norcross (Ed.), *Psychotherapy relationships that work: Therapist*

- contributions and responsiveness to patients* (pp. 235-254). New York, NY: Oxford University Press.
- Sakinofsky, I. (1979). Evaluating the competence of psychotherapists. *Canadian Journal of Psychiatry, 24*, 193-205.
- Samstag, L.W. (2006). The working alliance in psychotherapy: An overview of the invited papers in the special section. *Psychotherapy: Theory, Research, Practice, Training, 43*, 300-307.
- Samstag, L.W., Muran, C.J., Wachtel, P.L., Slade, A., Safran, J.D., & Winston, A. (2008). A comparison of working alliance, interpersonal behavior, and narrative coherency among three psychotherapy outcome conditions. *American Journal of Psychotherapy, 62*, 165-194.
- Schmidt, F.L., & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.
- Sexton, H., Littauer, H., Sexton, A., & Tommeras, E. (2005). Building an alliance: Early therapy process and the client-therapist connection. *Psychotherapy Research, 15*, 103-116.
- Shapiro, D.A., Barkham, M., Rees, A., Hardy, G., Reynolds, S., & Startup, M. (1994). Effects of treatment duration and severity of depression on the effectiveness of cognitive behavioral and psychodynamic-interpersonal psychotherapy. *Journal of Consulting and Clinical Psychology, 62*, 522-534.

- Shapiro, D.A., Firth-Cozens, J., & Stiles, W.B. (1989). The question of therapists' differential effectiveness: A Sheffield Psychotherapy Project addendum. *British Journal of Psychiatry*, 154, 383-385.
- Shapiro, D.A., & Shapiro, D. (1982). Meta-analysis of comparative therapy outcome studies: A replication and refinement. *Psychological Bulletin*, 92, 581-604.
- Shen-Miller, D.S., Grus, C.L., Van Sickle, K.S., Schwartz-Mette, R., Cage, E.A., Elman, N.S., Jacobs, S.C., & Kaslow, N.J. (2011). Trainees' experiences with peers having competence problems: A national survey. *Training and Education in Professional Psychology*, 5, 112-121.
- Shlien, J. (1997). Empathy in psychotherapy: A vital mechanism? Yes. Therapists conceit? All too often. By itself enough? No. In A.C. Bohart & L.S. Greenberg (Eds.), *Empathy reconsidered: New directions in psychotherapy* (pp. 63-80). Washington DC: American Psychological Association.
- Shrout, P.E., & Fleiss, J.L. (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86, 420-428.
- Sloane, R.B., Staples, F.R., Cristol, A.H., Yorkston, N.J., & Whipple, K. (1975). *Short-term analytically oriented psychotherapy vs. behavior therapy*. Cambridge, MA: Harvard University Press.
- Spruill, J., Rozensky, R.H., Stigall, T.T., Vasquez, M., Bingham, R.P., & Olvey, C.D. (2004). Becoming a competent clinician: Basic competencies in intervention. *Journal of Clinical Psychology*, 60, 741-754.
- Stahl J.V., & Hill, C.E. (2008). A comparison of four methods for assessing natural helping ability. *Journal of Community Psychology*, 36, 289-298.

- Stein, D.M., & Lambert, M.J. (1995). Graduate training in psychotherapy: Are therapy outcomes enhanced? *Journal of Consulting and Clinical Psychology, 63*, 182-196.
- Sternberg, R.J., & Williams, W.M. (1993). Does the Graduate Record Examination predict meaningful success in the graduate training of psychologists: A case study. *American Psychologist, 52*, 630-641.
- Stiles, W.B., Shapiro, D.A., & Elliott, R.K. (1986). Are all psychotherapies equivalent? *American Psychologist, 41*, 165-180.
- Strauss, J.L., Hayes, A.M., Johnson, S.L., Newman, C.F., Brown, G.K., Barber, J.P., Laurenceau, J.P., Beck, A.T. (2006). Early alliance, alliance ruptures, and symptom change in a nonrandomized trial of cognitive therapy for avoidant and obsessive-compulsive personality disorders. *Journal of Consulting and Clinical Psychology, 74*, 337-345.
- Stricker, G., & Callan, J.E. (1987). Attitudes and aptitudes of a professional psychologist. In E.F. Bourg et al. (Eds.), *Standards and evaluation in the education and training of professional psychologists* (pp. 129-140).
- Strong, S.R. (1968). Counseling: An interpersonal influence process. *Journal of Counseling Psychology, 17*, 388-399.
- Strupp, H.H. (1958). The psychotherapist's contribution to the treatment process. *Behavioral Science, 3*, 34-67.
- Strupp, H.H. (1973). On the basic ingredients of psychotherapy. *Journal of Consulting and Clinical Psychology, 41*, 1-8.

- Strupp, H.H. (1977). A reformulation of the dynamics of the therapist's contribution. In A.S. Gurman & A.M. Razin (Eds.), *Effective psychotherapy: A handbook of research* (pp. 1-22). Elmsford, NY: Pergamon Press.
- Strupp, H.H. (1995). The psychotherapist's skills revisited. *Clinical Psychology: Science and Practice*, 2, 70-74.
- Strupp, H.H. (1998). The Vanderbilt I study revisited. *Psychotherapy Research*, 8, 17-29.
- Strupp, H.H., & Binder, J.L. (1984). *Psychotherapy in a new key: A guide to time-limited dynamic psychotherapy*. New York, NY: Basic Books.
- Strupp, H.H., & Hadley, S.W. (1979). Specific versus nonspecific factors in psychotherapy: A controlled study of outcome. *Archives of General Psychiatry*, 36, 1125-1136.
- Suh, C.S., Strupp, H.H., & O'Malley, S.S. (1986). The Vanderbilt process measures: The Psychotherapy Process Scale (VPPS) and the Negative Indicators Scale (VNIS). In L.S. Greenberg, & W.M. Pinsoff (Eds.), *The psychotherapeutic process: A research handbook* (pp. 285-324). New York, NY: The Guilford Press.
- Tasca, G.A., & McMullen, L.M. (1992). Interpersonal complementarity and antitheses within a stage model of psychotherapy. *Psychotherapy*, 29, 515-523.
- Task Force on Promotion and Dissemination of Psychological Procedures. (1995). Training and dissemination of empirically validated psychological treatments: Report and recommendations. *The Clinical Psychologist*, 48, 2-23.
- Tepper, D.T., & Haase, R.F. (1978). Verbal and nonverbal communication of facilitative conditions. *Journal of Counseling Psychology*, 25, 35-44.

- Teyber, E., & McClure, F. (2000). Therapist variables. In C.R. Snyder & R.E. Ingram (Eds.), *Handbook of psychological change: Psychotherapy processes and practices for the 21st century* (pp. 62-87). New York, NY: John Wiley & Sons.
- Thompson, B., & Hill, C. (1993). Client perceptions of therapist competence. *Psychotherapy Research, 3*, 124-130.
- Timulak, L. (2007). Identifying core categories of client-identified impact of helpful events in psychotherapy: A qualitative meta-analysis. *Psychotherapy Research, 17*, 305-314.
- Trepka, C., Rees, A., Shapiro, D.A., Hardy, G.E., & Barkham, M. (2004). Therapist competence and outcome of cognitive therapy for depression. *Cognitive Therapy and Research, 28*, 143-157.
- Truax, C.B., & Carkhuff, R. (1967). *Toward effective counseling and psychotherapy: Training and practice*. Chicago, IL: Aldine.
- Truax, C.B., & Mitchell, K.M. (1971). Research on certain therapist interpersonal skills in relation to process and outcome. In A.E. Bergin & S.L. Garfield (Eds.), *Handbook of psychotherapy and behavior change: An empirical analysis* (pp. 299-344). New York, NY: John Wiley & Sons.
- Umphress, V.J., Lambert, M.J., Smart, D.W., Barlow, S.H., Clouse, G. (1997). Concurrent and construct validity of the Outcome Questionnaire. *Journal of Psychoeducational Assessment, 15*, 40-55.
- Vacha-Haase, T., Davenport, D.S., & Kerewsky, S.D. (2004). Problematic students: Gatekeeping practices of academic professional psychology programs. *Professional Psychology: Research and Practice, 35*, 115-122.

- Vanaerschot, G., & Lietaer, G. (2007). Therapeutic ingredients in helping session episodes with observer-rated low and high empathic attunement: A content analysis of client and therapist postsession perceptions in three cases. *Psychotherapy Research, 17*, 329-342.
- Vermeersch, D.A., Lambert, M.J., & Burlingame, G.M. (2000). Outcome Questionnaire: Item sensitivity to change. *Journal of Personality Assessment, 74*, 242-261.
- Vermeersch, D.A., Whipple, J.L., Lambert, M.J., Hawkins, E.J., Burchfield, C.M., & Okiishi, J.C. (2004). Outcome Questionnaire: Is it sensitive to changes in counseling center clients? *Journal of Counseling Psychology, 51*, 38-49.
- Vivino, B.L., Thompson, B.J., Hill, C.E., & Ladany, N. (2009). Compassion in psychotherapy: The perspective of therapists nominated as compassionate. *Psychotherapy Research, 19*, 157-171.
- von der Lippe, A.L., Monsen, J.T., Ronnestad, M.H., & Eilertsen, D.E. (2008). Treatment failure in psychotherapy: The pull of hostility. *Psychotherapy Research, 18*, 420-432.
- Wampold, B.E. (2001). *The great psychotherapy debate: Models, methods, and findings*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Wampold, B.E. (2007). Psychotherapy: The humanistic (and effective) treatment. *American Psychologist, 62*, 857-873.
- Wampold B.E., & Bhati, K.S. (2004). Attending to the omissions: A historical examination of evidence-based practice movements. *Professional Psychology: Research and Practice, 35*, 563-570.

- Wampold, B.E., & Brown, G.S. (2005). Estimating variability in outcomes attributable to therapists: A naturalistic study of outcomes in managed care. *Journal of Consulting and Clinical Psychology, 73*, 914-923.
- Wampold, B.E., Imel, Z.E., & Minami, T. (2007). The placebo effect: “Relatively large” and “robust” enough to survive another assault. *Journal of Clinical Psychology, 63*, 401-403.
- Wampold, B.E., Minami, T., Tierney, S.C., Baskin, T.W., & Bhati, K.S. (2005). The placebo is powerful: Estimating placebo effects in medicine and psychotherapy from randomized clinical trials. *Journal of Clinical Psychology, 61*, 835-854.
- Wampold, B.E., Mondin, G.W., Moody, M., Stich, F., Benson, K., & Ahn, H. (1997). A meta-analysis of outcome studies comparing bona fide psychotherapies: Empirically, “all must have prizes.” *Psychological Bulletin, 122*, 203-215.
- Watson, J.C. (2002). Re-visioning empathy. In D.J. Cain, & J. Seeman (Eds.), *Humanistic psychotherapies: Handbook of research and practice* (pp. 445-472). Washington DC: American Psychological Association.
- Willingham, W.W. (1974). Predicting success in graduate education. *Science, 183*, 273-278.
- Wilson, G.T., & Evans, I.M. (1977). The therapist-client relationship in behavior therapy. In A.S. Gurman & A.M. Razin (Eds.), *Effective psychotherapy: A handbook of research* (pp. 544-565). Elmsford, NY: Pergamon Press.
- Yalom, I.D., & Lieberman, M.A. (1971). A study of encounter group casualties. *Archives of General Psychiatry, 47*, 427-439.

Zuroff, D.C., & Blatt, S.J. (2006). The therapeutic relationship in the brief treatment of depression: Contributions to clinical improvement and enhanced adaptive capacities. *Journal of Consulting and Clinical Psychology, 74*, 130-140.

VITA

Jacqueline P. Camp

Degrees:

Bachelor of Arts, Psychology, 2004
University of Nevada, Las Vegas

Master of Arts, Psychology, 2007
University of Nevada, Las Vegas

Special Honors and Awards:

2006 Graduate Research Training Scholarship (GREAT) \$3,333

Publications:

Eno Loudon, J., Skeem, J., Camp, J., Peterson, J., & Vidal, S. (2010). Supervision practices in special mental health probation: What happens in officer-probationer meetings? *Law and Human Behavior*, online first, 23 Dec. 2010.

Kennealy, P., Skeem, J., Walters, G., & Camp, J. (2010). Do core interpersonal and affective traits of PCL-R Psychopathy interact with antisocial behavior and disinhibition to predict violence? *Psychological Assessment*, 22, 569-580.

Vidal, S., Skeem, J., & Camp, J. (2010). Emotional intelligence: Painting different paths for low-anxious and high-anxious psychopathic variants. *Law & Human Behavior*, 34, 150-163.

Eno Loudon, J., Skeem, J., Camp, J., & Christensen, E. (2008). Supervising probationers with mental disorders: How do agencies respond to violations? *Criminal Justice & Behavior*, 35, 832-847.

Skeem, J., Eno-Loudon, J., Polaschek, D., & Camp, J. (2007). Assessing relationship quality in mandated community treatment: Blending care with control. *Psychological Assessment*, 19, 397-410.

Skeem, J., Edens, J., Camp, J., & Colwell, H. (2004). Do racial groups differ in levels of psychopathy? A meta-analysis. *Law & Human Behavior*, 28, 505-527.

Dissertation Title: Relationship competence: Can trainee interpersonal skills be measured reliably and do they predict clinical effectiveness?

Dissertation Examination Committee:

Co-Chair, Christopher Heavey, Ph.D.

Co-Chair, Michelle Carro, Ph.D.

Committee Member, Marta Meana, Ph.D.

Committee Member, Kimberly Barchard, Ph.D.

Graduate College Representative, Stephen Fife, Ph.D.