The Relationship between Therapist Empathy, the Working Alliance, and Therapy

Outcome: A Test of a Partial Mediation Model

A dissertation presented to

the faculty of

the College of Arts and Sciences of Ohio University

In partial fulfillment

of the requirements for the degree

Doctor of Philosophy

Edgar H. Wing, Jr.

March 2010

This dissertation titled

The Relationship between Therapist Empathy, the Working Alliance, and Therapy

Outcome: A Test of a Partial Mediation Model

by

EDGAR H. WING, JR.

has been approved for

the Department of Psychology

and the College of Arts and Sciences by

_____

Timothy M. Anderson

Professor of Psychology

_____

Benjamin M. Ogles

Dean, College of Arts and Sciences

ABSTRACT

WING, EDGAR H., Ph.D., March 2010, Psychology

<u>The Relationship between Therapist Empathy, the Working Alliance, and Therapy</u>

<u>Outcome: A Test of a Partial Mediation Model</u> (58 pp.)

Director of Dissertation: Timothy M. Anderson

Therapist empathy and the working alliance are two of the most potent predictors of therapy outcome (Horvath & Bedi, 2002; Martin, Garske, & Davis, 2000; Bohart, Elliott, Greenberg & Watson, 2002); yet the relationship among these variables is poorly understood. Using a random subset of 30 clients from the Vanderbilt 2 psychotherapy study (Strupp, 1993; Henry, Strupp, Butler, Schacht and Binder, 1993; and Henry, Schacht, Strupp, Butler and Binder, 1993) sample, the current study tested a mediational model in which empathy was predicted to have an influence on outcome indirectly through the alliance and directly, independent of the alliance. Bootstrap and Ordinary Least Squares regression analysis results supported the mediational role of the alliance in the relationship between empathy and therapy outcome; however, no direct effect of empathy on outcome independent of the alliance was found.

Approved: _____

Timothy M. Anderson

Professor of Psychology

ACKNOWLEDGMENTS

I would like to express my appreciation to my committee members—Tim, Ben, John, Ken, and Pat—for their time, effort, patience, and guidance in this project. I wish to thank my advisor, Tim Anderson, especially for his continued, good-natured support through this long process. I would also like to thank my father, Edgar Wing, Sr., for his example of the pursuit of knowledge for its own sake and my brother, Tom, for his inspiring example of living life to the fullest. Finally, I would like to thank my wife, Aubri Deahl Wing, for her love and support without which this project might never have seen its completion.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1: INTRODUCTION

Therapeutic empathy and the working alliance between therapist and client are two of the most robust predictors of psychotherapy outcome yet identified by psychotherapy researchers (e.g., Horvath & Bedi, 2002; Martin, Garske, & Davis, 2000; Bohart, Elliott, Greenberg & Watson, 2002); yet the relationship between these process variables is poorly understood. Given the demonstrated potency of these interpersonal process variables, further elucidation of their relationship to each other and to psychotherapy outcomes could lead to improvements in psychotherapy training and success. The purpose of the current study is to test a model of the relationship among therapist empathy, the working alliance, and psychotherapy outcome. Specifically, it is argued that therapist empathy has its influence on therapy outcome partially through its influence on the development and maintenance of the working alliance, but also directly, via processes independent of the alliance.

Theorists from diverse theoretical perspectives have suggested a strong relationship between empathy and the alliance. Psychodynamic theorists (e.g., Freud, 1913; Zetzel, 1956; Greenson, 1960) have tended to emphasize the role of an empathic therapist stance in developing an affective bond with the client, often conceptualized as part of a positive transferential process. Some modern humanistic theorists (e.g., Elliott, Watson, Goldman & Greenberg, 2004) have emphasized the role of therapist empathy in creating a safe environment in which the client is freed to focus on the work of therapy rather than the relationship itself, while Watson (2002) noted the importance of therapist empathy in the process of negotiating agreement on the tasks and goals of therapy. Coming from a cognitive-behavioral orientation, Raue & Goldfried (1994) also

acknowledge an important role for therapist empathy in the development of a therapeutic

bond, which is then used to encourage and support desired behavior change. In addition

to its role in the development of a working alliance, a number of writers have emphasized

the role of therapist empathy in the detection and resolution of alliance ruptures (e.g.,

Burns, 1989; Burns & Auerbach, 1996; Safran & Segal, 1990; Linehan, 1993).

While researchers and theorists from a wide range of theoretical perspectives have

acknowledged a role for therapist empathy in the development and maintenance of the

alliance, other functions of empathy in psychotherapy, independent of its role in the

alliance, have also been discussed. For instance, Watson (2002) argues that "[e]mpathic

responses can assist clients in deconstructing their world views so that they can become

aware of the subjectivity of their perceptions" which "assists them in being more

hypothetical in their formulations of events so as to increase their range of action" (p.

462). Rogers (1975) noted that therapist empathy facilitated clients' discovery of new

aspects of themselves and that this contributed to a change in their self-concept and,

consequently, behavior. Watson (2002) argues that empathic therapist responses help

clients to regulate affect by helping clients to access and process their emotions. Through

the process of reflexively accessing and symbolizing their emotions, clients come to

better understand and more adaptively use them to achieve valued goals. Finally, both

Rogers (1975) and Watson (2002) argue that the experience of being listened to

empathically teaches clients to be more accepting and nurturing of themselves.

Taken together, these theoretical accounts imply a process in which therapist

empathy has its influence on therapy success both via its influence on the development

and maintenance of a therapeutic alliance and via other processes relatively independent

of the alliance. In other words, this model is one of partial mediation, in which therapist

empathy's influence on outcome is partially mediated by its influence on the therapeutic alliance and partially independent of the alliance. Despite the theoretical links between empathy and the working alliance, empirical studies of these hypothesized relationships are relatively scarce. The few existing studies provide direct or indirect evidence for aspects of the partial mediation model, addressing questions of whether (1) there is a relationship between empathy and alliance, (2) whether empathy and alliance are empirically separable constructs, and (3) whether empathy and alliance have separable influences on psychotherapy outcome.

The Empirical Relationship between Empathy and Alliance

A few studies have examined the relationship between empathy and the therapeutic alliance, in each case finding a significant moderate to large correlation between ratings of the two constructs. Salvio, Beutler, Wood, and Engle (1992) had clients (N = 46) rate empathy and alliance at the termination of a 20-session course of therapy using the BLRI Empathy subscale and the WAI. In that study, BLRI Empathy scale scores correlated with WAI scales 0.65 to 0.85. In a study by Horvath (1981) both empathy and alliance were measured from the client's (N = 29) perspective using the BLRI Empathy subscale and the WAI, respectively, after session 3. Empathy correlated with the goal, task, and bond subscales of the WAI 0.63, 0.63, and 0.83, respectively. In a similar study, Mosely (1983; N = 31) found session 3 correlations between client-rated BLRI Empathy and WAI goal, task, and bond scales to be 0.70, 0.70, and 0.76, respectively. In a study of relationship variables in the NIMH Treatment of Depression Collaborative Research Project, Zuroff, Blatt, Sotsky, Krupnick, Martin, Sanislow, and Simmens (2000; N = 149) examined the relationship between empathy and alliance using a composite of the highly intercorrelated BLRI Empathy, Level of Regard, and

Congruence subscales and the Patient Contribution subscale of the Vanderbilt Therapeutic Alliance Scale, tapping collaborative and other client contribution aspects of alliance. They found a small, but significant, correlation ($r = 0.22$) between session 2 empathy and session 3 alliance and a similar relationship at the end of therapy ($r = 0.25$). Watson & Geller (2005) reported a correlation of 0.72 between the mean BLRI full scale scores from sessions 9 and 12 and mean WAI full scale scores from each of 16 sessions of therapy. Given the fact that all of these studies measured only clients' perceptions of therapist empathy, these findings must be considered only indirect evidence of the relationship between actual therapist empathic behavior and the therapeutic alliance. More direct evidence of the relationship between therapist empathy and alliance would come from a study measuring actual therapist behavior using objective observers. Observer ratings would presumably be relatively free of a number of the potentially contaminating influences in clients' perceptions (e.g., halo effects, transference, mood; Duan & Hill, 1996).

Are Empathy and Alliance Empirically Separable Constructs?

The second question of interest to the current study is whether therapist empathy and alliance are empirically separable constructs. This question arises out of the contention by some (Salvio et al., 1992; McCabe & Priebe, 2004; Waddington, 2002) that high correlations between measures of therapist empathy and measures of the alliance imply that they are each part of the same underlying process and, therefore, empirically inseparable. On this question the existing empirical evidence is equivocal.

Salvio et al. (1992) conducted a principle axis factor analysis with Varimax rotation ($N = 32$) of the 3 WAI subscale scores from sessions 5, 10, 15, and 20 (termination) and the 4 BLRI subscale scores from session 20. All subscales from all

sessions showed high loadings on a single factor. The BLRI Empathy subscale loaded 0.72 on this factor. This, in combination with the high intercorrelations of the WAI and BLRI subscales noted above, has led some (e.g., Salvio et al., 1992; McCabe & Priebe, 2004; Waddington, 2002) to conclude that therapist empathy and the working alliance are overlapping, if not essentially identical, constructs, or, at least, are difficult to discriminate.

Several limitations to the Salvio et al. (1992) study prohibit drawing such a general conclusion, however. First, empathy and alliance were both measured from the client's perspective, allowing for an unknowable amount of common method variance to inflate correlations. Contributing further to this possibility is the fact that empathy was measured only at the end of a 20-session course of therapy, likely increasing the probability of halo effects. For instance, it seems likely that clients were judging their therapists empathy at this late stage based to a great extent on their accumulated experience with them over time, allowing them to pick and choose from among their varied experiences according to their general feelings about their therapist, experiences likely correlated with the strength of the therapeutic alliance. In this regard it is interesting to note that the factor loadings for the WAI scales are invariably higher at the latter 2 administrations (at sessions 15 and 20)—and invariably highest at termination (session 20)—than at the earlier sessions (5 and 10). This suggests the possibility that clients' ratings of these constructs became less differentiated over time. It is possible that this lack of differentiation reflects an actual change in the relationship among these 3 alliance components, such that they became more highly correlated with each other over the course of therapy, especially after therapy ended. However, it seems at least equally plausible that the increasing correspondence among clients' perceptions of WAI alliance

components reflects clients' general feelings about their therapists and/or the outcome of their therapy (e.g., Kazdin, 2007). Such an effect could reasonably be assumed to affect the correlations among the BLRI scales—rated only at termination—and WAI scales rated at that time. This would have the effect of inflating correlations between BLRI and WAI scales, leading to the overestimation of the overlap in these constructs.

In terms of understanding how therapists can contribute to the alliance, the Salvio et al. study is limited by its use of a client-rated measure of empathy. While clients' experience of empathy has been found to be a potent predictor of outcome (Bohart et al., 2002), this does not directly tell us about therapists' behavior, especially considering the typically low correlations between observer and client ratings of therapist empathy (Bohart & Greenberg, 1997). As noted by Duan and Hill (1996), client-rated measures of therapist empathy are limited by a host of possible errors in human perception that have nothing to do with actual therapist behavior. For instance, client ratings may be a reflection of client mood, cognitive/affective distortions (e.g., transference), or halo effects based on therapy outcome and other processes.

Another caveat in interpreting Salvio et al.'s study is that the subject-item ratio for their factor analysis was only 2:1, lower than typically recommended (Guadagnoli & Velicer, 1988). Finally, even high correlations between empathy and alliance at a given time do not necessarily imply that they are aspects of the same construct. For instance, if empathy is a necessary and/or sufficient condition for alliance development, then high correlations between the two are to be expected. What seems safe to conclude from this study is that clients' perceptions of therapist empathy and the working alliance tend to be strongly correlated later in therapy.

Further support for the empirical inseparability comes from the high correlations (see above) between BLRI Empathy subscale scores and WAI subscale scores found by Horvath (1981) and Moseley (1983), who had actually been attempting in their studies to provide evidence of discriminant validity of the WAI. In each of these studies, both empathy and alliance were measured from the client's perspective after session 3. As with Salvio et al.'s (1992) study discussed above, the fact that Horvath (1981) and Moseley (1983) measured both empathy and alliance from the clients' perspectives and at the same point in time creates the possibility for halo effects and other common-method problems (e.g., response-style and mood effects), thus possibly inflating correlations. As also noted above, measuring empathy from only the client's perspective does not directly tap into therapist-offered empathy. A strength of the Horvath (1981) and Moseley (1983) studies is that empathy and alliance are measured relatively early in the therapies (session 3), providing indirect evidence of the possible influence of empathy in the development of the alliance.

When Zuroff et al. (2000), in the study described above, simultaneously regressed a composite outcome measure on their session 2 empathy-like BLRI composite and the session 3 alliance-like Patient Contribution subscale, the two predictors were independently related to outcome. The BLRI composite was positively related to outcome, beta = 0.18, sr2 = .032, $p < 0.05$, as was the Patient Contribution factor of the VTAS, beta = 0.26, sr2 = .065, $p < 0.01$. These results provide some evidence for unique contributions of empathy and alliance to therapy outcome and, therefore, evidence for the separability of the two constructs. These findings are inconsistent with Watson and Geller (2005) and Horvath (1981) who found a BLRI composite and BLRI Empathy,

respectively, to no longer significantly predict outcome after the WAI was entered into the regression model.

A number of limitations make the results of Zuroff et al. (2000) difficult to interpret for present purposes. First, while their BLRI composite seems to represent an empirically coherent construct, with the BLRI Empathy scale presumably correlating highly with it, it cannot be considered an unequivocal measure of empathy. Second, while the empirically derived Patient Contribution factor of the VTAS used in this study has some face validity as a measure of alliance, it also apparently contains other client contributions somewhat outside a coherent collaboration construct and has not been validated as a measure of alliance. Finally, the empathy-like construct, as in other studies reviewed here, was assessed from the perspective of the client, providing only indirect evidence, at best, for the influence of therapist behavior in these processes. Despite these limitations, this study provides good tentative evidence for a model in which therapist empathy impacts outcome partially through its influence on the development and maintenance of the alliance and partially through other, separate pathways.

Further evidence of the separable influences of empathy and alliance on outcome comes from a study by Gaston, Marmar, Gallagher, and Thompson (1991). In this study, older depressed adults participating in one of 3 treatment conditions completed the CALPAS after sessions 5, 10, and 15. Gaston et al. then examined the unique relationships between each of the 5 subscales comprising the CALPAS and outcome for each of the treatment samples and for each administration of the CALPAS. Of present interest is the finding that Therapist Understanding contributed unique variance to outcome over and above the other components of alliance, with partial correlations ranging from +0.11 to -0.47, with 6 of 9 correlations in the expected negative direction.

Small samples sizes in this study precluded statistical tests of significance. While

promising, these results provide only indirect evidence of the unique influence of

therapist empathy (controlling for alliance) on outcome. In addition to problems with

small sample sizes, the empathy-like ratings were provided by clients, creating potential

problems of common method variance and halo effects with the client-rated alliance

constructs. Finally, and perhaps most importantly, the Therapist Understanding

component of the CALPAS assesses constructs in addition to therapist empathy, making

it difficult to interpret for current purposes.

<div align="center">Empathy, Alliance, and Outcome</div>

A final question relevant to the current study is whether empathy has its

demonstrated influence on therapy outcome (Bohart et al., 2002) completely through its

influence on the alliance or through multiple pathways, as suggested by theory. Evidence

relevant to this question comes from the mediation analyses of Watson and Gellar (2005)

who attempted to answer this question directly, though with equivocal results, and from

Zuroff et al. (2000) who, while not explicitly testing for mediation, reported results

consistent with a partial mediation effect. While Watson and Geller found the BLRI

scores no longer to be significant predictors of outcome after the influence of alliance

was partialled out, Zuroff et al. found empathy and alliance to have unique influences on

outcome, suggesting the possibility of an only partial mediation effect. Limitations of

these studies have already been discussed above. What is required to answer the question

of the mediating influence of alliance in the relationship between empathy and outcome

is a study in which empathy is measured early in therapy using a relatively pure measure

of empathy and alliance is measured either concurrently with empathy or later in therapy.

The current study is designed to address the broad question of the empirical separability of therapist empathy and the working alliance as well as the more specific question of the nature of the relationship between these two constructs. First, the separability of empathy and alliance is tested by examining the pattern of relationships among empathy, alliance, and other psychotherapy process variables that are expected to show different relationships with empathy and alliance, respectively. Further, this study was designed to test a model of empathy-alliance-outcome relationship implied by the theoretical literature. The model, as formulated here, states that therapist empathy has a positive impact on outcome, both through an impact on the alliance (which, in turn, impacts outcome) and through other processes relatively independent of the working alliance.

CHAPTER 2: METHOD

Data for this study were based on the Vanderbilt II psychotherapy project. This study was designed to examine the effect of therapist training on psychotherapy outcome, and is reported in Strupp (1993), Henry, Strupp, Butler, Schacht and Binder (1993), and Henry, Schacht, Strupp, Butler and Binder (1993).

Clients

The clients from the current study were selected from the Vanderbilt II video archive, a study that originally included 64 clients. Information about selection and procedures from this study can be found in Appendix A and in Henry et al. (1993). Out of the original sample of 64, 54 provided sufficient data for some analyses, the sample for other analyses being restricted by further sampling, as described next.

Alliance data for the current study (i.e., WAI and GAR scores) were collected in a study subsequent (Wang & Anderson, unpublished study) using a randomly selected subsample of 30 of these 54 cases; therefore, all analyses using alliance data are based on this smaller sample.

Equivalence of samples was tested by comparing the sample of 54 with the subsample randomly chosen by Wang and Anderson on a number of demographic and study-relevant variables. Without exception, no significant difference was found between the 2 samples (N=54 vs. N=30, respectively) in terms of mean client age (42 vs. 41), percentage of female clients (78 vs. 77), percentage of white clients (98 vs. 97), percentage of clients with 2 or more years of college (72 vs. 70), percentage of clients married (46.3 vs. 43.3) , number of sessions completed (21.2 vs. 21.9), (MEE total score;

5.1 vs. 5), Patient Participation (z = -.06 vs. -.05) or Therapist Warmth & Friendliness (z = .07 vs. -.31) ratings, symptom improvement from intake to termination (residualized GSI scores; z = .00 vs. .16), percentage of clients with previous therapy experience (72.2 vs. 83.3), percentage of clients with at least one Axis I diagnosis (98.1 vs. 96.7), or percentage of clients with at least one Axis I diagnosis (both 100%).

Therapists

Therapists in the study were 8 licensed clinical psychologists and 8 psychiatrists who had been recommended by previous supervisors as caring, empathic clinicians. All had at least 2 years full-time post-degree experience, with an average experience level of 5.6 years, and previous training and supervision in psychodynamic principles. Ten of the therapists were men, 6 were women; all were Caucasian.

Therapy

Approximately half the clients in the original Vanderbilt II study received standard treatment from mostly psychodynamically oriented therapists, the other half receiving Time Limited Dynamic Psychotherapy, which was designed to address relationship problems.

Measures

The data for the current study come from a series of studies, all using the Vanderbilt II Psychotherapy Project sample described above. The SCL-R-90 and Vanderbilt Psychotherapy Process Scales data, as well as information about client and therapist characteristics, come from the original Vanderbilt II study. The 2 alliance measures—the Global Alliance Rating and the Working Alliance Inventory (manualized

observer-rated version) come from an unpublished study by Wang & Anderson using the Vanderbilt II videotape archive. Data for the 2 empathy scales—the Measure of Expressed Empathy and the Accurate Empathy Scale—were collected as part of the current study using the Vanderbilt II videotape archive.

Measure of Expressed Empathy (MEE; Watson, 1999). This is an observer-rated measure of therapist-communicated empathy. The measure evaluates therapists' verbal and non-verbal behaviors, speech characteristics, and response modes. It consists of 10 items that are rated in terms of frequency on a nine-point Likert scale ranging from 0 (never) to 8 (all the time). Total scores range from 0 to 80. Interrater reliability for two raters was found to be $r = .87$ ($p < 0.01$; Steckley, P.L., 2006). Evidence of convergent validity of an earlier, 9-item version was evidenced by a significant and high correlation ($r = .66$) with a client-rated version of the BLRI Empathy Scale (Barrett-Lennard, 1962; Watson & Prosser, 2002). In a more recent study (Steckley, P.L., 2006), MEE scores averaged across sessions 3 and 15 correlated $r = .29$ ($p < 0.05$) with BLRI Empathy scores averaged across sessions 9 and 12. Internal consistency reliability (Cronbach's alpha) for the 10 MEE items for this study was .92.

In the current study, MEE ratings are based on videotapes of the middle 20 minutes of the third session of therapy. The author of the study provided all ratings and interrater reliability ($r = .81$, $p < 0.05$) was tested by correlating these ratings with those of a second rater on a random selection of 50% of the sample.

Truax-Carkhuff Accurate Empathy Scale (AES; Truax & Carkhuff, 1967). This scale was designed to measure Rogers' (1957) conception of accurate therapist empathy.

It consists of descriptions of 9 increasing levels of accuracy in empathic responding.

Scores range from 1 to 9, reflecting the level that best represents the responding of the

therapist to the client's communications during the segment rated. The scale has been

used widely in psychotherapy research, which has provided abundant evidence of its

interrater reliability and predictive validity with respect to a variety of therapy outcomes

(e.g., Truax & Carkhuff, 1967; Bohart et al., 2002). In the current study, raters provided

ratings of accurate therapist empathy based on the middle 20 minutes of the third session

of therapy. The author of the study provided all ratings and interrater reliability (r = .58,

p < 0.05) was tested by correlating these ratings with those of a second rater on a random

selection of 50% of the sample. Given the relatively low interrater reliability for this

measure, its use in the current study is restricted to peripheral analyses secondary to the

central hypotheses being tested.

   Working Alliance Inventory (Observer-rated version; WAI-OM; Wang &

Anderson). The WAI (Horvath & Greenberg, 1986) is a 36-item self-report measure

based on Bordin's (1979) transtheoretical formulation of the therapeutic alliance. The

WAI is comprised of three empirically derived subscales: Task Agreement, Goal

Agreement, and Bond Development which each yield individual subscale scores in

addition to one combined overall index score. Each subscale contains twelve items

scored on a seven-point Likert scale.

   High internal consistencies have been reported for the complete WAI (e.g.,

coefficient alpha of .96 for WAI-C and .95 for the WAI-T; Tichenor & Hill, 1989) as

well as individual subscales (e.g, alpha coefficients range from .85 to .88 for subscales of

the WAI-C and from .68 to .87 for subscales of the WAI-T; Horvath & Greenberg, 1989).

In addition, multiple findings support the convergent and discriminant validity of the

WAI Goal and Task scales, but evidence regarding the convergent validity of the Bond

scale is equivocal (Greenberg & Pinsof, 1986; Horvath & Greenberg, 1989). However,

some studies have failed to report significant correlations between WAI-T and WAI-C

scores, suggesting that measurement of the alliance from different perspectives can yield

different results (Tichenor & Hill, 1989). Horvath and Greenberg (1989) reported the

WAI-C to demonstrate strong predictive validity given the scale's significant correlations

with counseling outcome measures such as the Patient Post-therapy Questionnaire (CPQ;

Strupp, Wallach, & Wogan, 1964), State-Trait Anxiety Inventory (STAI; Spielberger,

Gorsuch, & Lushene, 1970) and the Scale of Indecision (SI; Osipow, Carney, & Barak,

1976).

The current study made use of an observer-rated version of the WAI. Ratings

were provided by independent observers in a previous validation study of an observer

rating manual developed by Wang & Anderson. Four coders (2 undergraduate and 2

graduate psychology students) rated segments (50% 15-minutes in length; 50% 45-

minutes in length) of videotapes of session 3. The intraclass correlation coefficient for

the 4 raters was relatively low (alpha = .289). Given the low interrater reliability for this

measure, its use in the current study is restricted to peripheral analyses secondary to the

central hypotheses being tested.

Global Alliance Rating (GAR; Wang & Anderson). The GAR is a global measure

of the alliance between therapist and client evidenced in a segment of observed therapy.

Four coders (2 undergraduate and 2 graduate psychology students) rated segments (50% 15-minutes in length; 50% 45-minutes in length) of videotapes of session 3 as part of a previous study (Wang & Anderson, unpublished) using the Vanderbilt 2 psychotherapy video archive. Interrater reliability for the 4 raters was alpha = .694. Convergent validity in this initial study was demonstrated by moderate to high correlations theoretically related variables, including a manualized observer-rated version of the WAI ($r = 0.87$, $p < 0.001$), the VPPS Patient Participation scale ($r = 0.5$, $p < 0.01$), and the VPPS Patient Hostility scale ($r = -0.45$, $p < 0.05$). Discriminant validity was supported by its lack of relationship with the VPPS Therapist Warmth and Friendliness scale ($r = 0.13$, n.s.). Predictive validity of the GAR is demonstrated by its significant moderate correlation with residualized change scores of the Global Severity Index (negative scores indicate positive change) in the current data ($r = -0.45$, $p < 0.05$).

 Vanderbilt Psychotherapy Process Scale (VPPS; O'Malley, Sue & Strupp, 1983). The VPPS consists of 80 items grouped to provide 3 global ratings (quality of client/therapist relationship; productiveness of the session; and how well the client is getting along); 40 ratings of client activities or characteristics; and 37 ratings of therapist activities or characteristics. Each rating is made on a scale from (1) not at all to (5) great deal. A factor analysis of VPPS items conducted by O'Malley et al. (1983) yielded 8 factors: Patient Participation; Patient Hostility; Patient Psychic Distress; Patient Exploration; Patient Dependency; Therapist Exploration; Therapist Warmth and Friendliness; and Negative Therapist Attitude.

Symptom Checklist-90-Revised (SCL-90-R (Derogatis, 1983). The SCL-90-R is a 90-item psychological symptom self-report measure. Items represent problems that have distressed patients during the past 7 days. Each item is rated on a 5-point scale of distress with a 0 response indicating "not-at-all" and a 4 reponse indicating "extremely". Items comprise 9 primary symptom dimensions: somatization, obsessive-compulsive, interpersonal sensitivity, depression, anxiety, hostility, phobic anxiety, paranoid ideation, psychoticism, and three global indices of distress: Global Severity Index (GSI), positive symptom distress index, and positive symptom total.

Reliability measures for the 9 primary symptom dimensions consist of internal consistency coefficient alphas ranging from .77 to .90 and test-retest reliability coefficients ranging from .78 to .90 for a one week interval. The SCL-90-R has been shown to reflect a high degree of convergent validity in studies comparing its symptoms scales to the MMPI clinical and content scales. Correlation coefficients ranged from .42 to .64 for the MMPI clinical scales and from .40 to .75 for the MMPI content scales (Derogatis, 1983).

CHAPTER 3: RESULTS

Zero-order relationships among Therapist Empathy, Alliance, and Related Constructs

Table 1 presents zero-order Pearson correlations among all process variables used in the current study. In further support of the convergent validity of the MEE, MEE scores were significantly positively correlated with the VPPS Therapist Warmth and Friendliness scale. As expected, the MEE was significantly correlated with the Global Alliance Rating (GAR).

Discriminating Therapist Empathy and Alliance

Table 1 also shows divergent patterns of relationships for Therapist Empathy and Alliance, supporting the hypothesis that empathy and alliance are separable constructs. As expected, alliance, unlike therapist empathy, did not correlate significantly with Therapist Warmth and Friendliness. The difference between these two correlation coefficients was only marginally significant, however $[Z(30) = 1.43, p \text{ (one-tailed)} = 0.08]$. As expected, empathy, unlike alliance, was not significantly related to the alliance-like variable Patient Participation. The difference between these two correlation coefficients was significant $[Z(30) = 1.87, p \text{ (one-tailed)} < 0.05]$.

Zero-order relationships between Therapist Empathy, Alliance, and Outcome

After GSI scores at termination were regressed on GSI scores at intake, the standardized residuals were used as Symptom Severity change scores as a measure of therapy outcome. Higher values on the GSI indicate greater severity of psychiatric symptomatology. Contrary to expectation, the MEE was not correlated significantly with GSI change $(r = -.05, \text{n.s.})$. As expected, there was a moderately strong negative

correlation between the GAR and GSI change (r = -.45, p < 0.05), indicating that a

stronger alliance at session 3 was predictive of greater decreases in symptom severity at

termination. The difference between these two correlation coefficients was significant

[Z(30) = 1.81, p (one-tailed) < 0.05]. While unexpected, the difference between the

relationships between empathy and outcome and alliance and outcome provides further

support for the separability of these two constructs.

Mediation Analyses

Having established significant relationships between therapist empathy and client-

therapist alliance and between client-therapist alliance and outcome, the hypothesis that

alliance mediates the relationship between therapist empathy and outcome was then

tested. There are a number of ways of testing mediation effects. Probably the most

commonly used among these is a hierarchical-regression-based method popularized by

Baron & Kenny (1986). Part of this method involves testing the significance of the

indirect effect of a proposed mediator using a test developed by Sobel (1982). Critical

reviews (e.g., Shrout & Bolger, 2002) and at least one Monte Carlo study (MacKinnon,

Lockwood, Hoffman, West & Sheets, 2002) have called this commonly-used method into

question, however. Most notable for present purposes is the finding that the Baron and

Kenny method is underpowered, making it inappropriate for smaller samples. Shrout and

Bolger (2002) recommended using an alternative, bootstrap method for testing both direct

and indirect effects in mediation analyses. In contrast to the more typical theoretical,

formula-based method, bootstrapping is a computer-based empirical method much akin

to those used in Monte Carlo studies. Bootstrapping involves the computer generation of

a large number (e.g., thousands) of samples from independent random sampling of the available sample. Confidence intervals are then generated from this pseudo-recreated sampling distribution. In addition to producing more accurate confidence intervals (when used in conjunction with a bias correction), Efron and Tibshirani (1993) found this method to be appropriate for samples as small as 20 to 80. Therefore, this bootstrapping method was the method used to test the significance of indirect and direct effects for all mediation hypotheses in this study.

In addition to a significant indirect effect, the logic of mediation requires that (1) there is a significant zero-order relationship between the predictor and the proposed mediator and (2) there is a significant zero-order relationship between the proposed mediator and the criterion (e.g., Shrout & Bolger, 2002). While Baron and Kenny's (1986) method calls additionally for the establishment of a significant zero-order relationship between the predictor and the criterion, others (e.g., Shrout & Bolger, 2002; Collins, Graham, & Flaherty, 1998; MacKinnon, Krull & Lockwood, 2000) have argued that this is not strictly necessary, especially when examining the relationship between distal predictor and criterion variables (e.g., early vs. late in therapy), as is this case in the current study.

Table 3 presents results of the mediation analysis using ordinary least squares regression and the bootstrap method described above. Five thousand bootstrap resamples were used to estimate the 95% confidence interval for the indirect effect. After listwise deletion of cases with missing data, the sample size for the mediation analysis was 28. In support of the mediation hypothesis, the indirect effect of therapist empathy on symptom

severity change through its influence on alliance was significant (B = -2.68, p < 0.05).

The negative sign of this effect indicates that empathy's influence on outcome via the

alliance was to reduce symptom severity, as would be expected. Controlling for the

number of completed sessions or whether the therapist-client dyad was part of the

treatment group (i.e., received TLDP) or the control group (i.e., received non-TLDP

therapy) did not change the results of the mediation analysis or any of its components.

Repeating the mediation analysis using an alternate empathy measure (AES) and/or an

alternate alliance measure (WAI-OM) yielded highly similar results, with significant

indirect effects for each combination of variables (see Tables 4, 5, and 6).

CHAPTER 4: DISCUSSION

This study was designed to elucidate the relationship between observable therapist empathic behavior and the alliance between therapist and client. Specifically, the study tested several hypotheses derived from the theoretical literature relating to these variables: (1) There is substantial relationship between observable therapist empathic behavior and alliance; (2) therapist empathy and the alliance are separable processes with independent relationships to other therapeutic processes and outcomes; (3) the influence of therapist empathy on therapy success occurs at least partially through its influence on the working alliance; and (4) therapist empathy has an influence on therapeutic success that is partially independent of its influence on the working alliance.

As expected, therapist empathy and alliance were found to be significantly and moderately correlated. While this replicates past studies, this is the first study to test a relationship between a relatively pure observer-rated measure of therapist empathy and a relatively pure observer-rated measure of the alliance. This finding is significant because previous studies have relied primarily upon clients' perceptions which are notoriously prone to distortions of actual therapist behavior (i.e., transference; Duan & Hill, 1996).

The results of this study provide support for the empirical separability of therapist empathy and the working alliance. Not only were empathy and alliance only modestly correlated, but they were differentially correlated with other therapy process variables in theoretically coherent ways. Alliance—but not therapist empathy—was related to Patient Participation, a construct reflecting the presence of such client behaviors as active participation in interactions with the therapist and taking the initiative in bringing up the

subjects that were talked about, as opposed to the client being withdrawn and inhibited (Suh et al., 1989). Patient Participation reflects the client's positive involvement in the work of therapy, which may have relatively little to do with therapist empathy, especially early in treatment when the client likely has come to therapy with his or her own motivation to work towards change. Therapist empathy—but not alliance—was related positively to Therapist Warmth and Friendliness, a measure of the therapist's warmth and emotional involvement with the client (Suh et al., 1989). This is a measure, then, reflecting an aspect of what the empathic therapist may bring to the therapeutic relationship, but an aspect which is not clearly central to the therapist and client working productively together. Therapist empathy—but not alliance—showed a negative relationship to Patient Psychic Distress. This suggests that therapist empathy may help reduce client distress, but that distress in itself does not necessarily impair the ability of therapists and clients to work productively together. To the extent that the reduction of distress is itself a goal of therapy, these results may reflect one way therapist empathy positively influences outcome independently of its relationship to the alliance. Taken together, these results call into question conclusions made by some researchers (e.g., Salvio et al., 1992) that empathy and alliance are aspects of the same underlying construct. Rather, these results provide support for the conceptualization of therapist empathy and the alliance as separable processes whose relationship to each other is only just beginning to be explored empirically.

Results provide preliminary support for the hypothesis that empathy's positive influence on outcome comes at least partially through its influence on the alliance. Given

that ratings of therapist's actual empathy behaviors were used in these analyses, this would suggest at least one way therapist's actions may influence therapy outcomes. Several limitations of this study, however, indicate caution in interpreting these results. First, while therapist empathy has been shown rather consistently to be positively related to therapy success (Bohart et al., 2002), no such relationship was found to be significant in the current study. While such a relationship is not required in testing a hypothesized indirect effect, the interpretation of the indirect effect becomes more difficult without it. This unexpected finding could be due to several different factors. Most prominent is the possible influence of the relatively unusual sample of clients and therapists used in the Vanderbilt 2 study. The clients in this study were selected for the presence of significant interpersonal difficulties. Indeed, as noted above, fully 67% of the original sample was diagnosed with at least one personality disorder. This is significant because Beutler, Crago, and Arizmendi (1986, p. 279) cite evidence that suggests that "patients who are highly sensitive, suspicious, poorly motivated, and reactive against authority…"–characteristics frequently found among those with the most common personality disorders—"…perform relatively poorly with therapists who are particularly empathic, involved, and accepting." Furthermore, therapists used in the Vanderbilt study were relatively highly experienced. This is significant because Bohart et al. (2002) found in their meta-analysis of the relationship between empathy and outcome a significant negative relationship between therapist experience and the magnitude of the relationship between empathy and outcome (effect level: $r = -0.24$; study level: $r = -0.43$). Clearly,

the issue of how client and therapist variables interact with empathy to influence—

positively or negatively—therapy outcomes is a fertile area for future research.

A second limitation of the current study concerns the difficulty of inferring causal

processes from correlational data. The logic of mediation requires the predictor variable

to represent a process causally prior to that of the proposed mediator; however, these

correlational data do not allow a claim of causal relationships. Furthermore, given that

empathy and alliance were measured at the same session, even the more modest claim of

temporal priority of the predictor cannot be made with any certainty. Future studies

should test the mediation hypothesis using empathy ratings from earlier points in therapy

than the alliance ratings, at least establishing temporal priority.

The mediation analysis did not support a unique relationship between empathy

and therapy outcome, that is, independent of the influence of the alliance. The mediation

results taken alone would suggest that empathy's influence on outcome—at least

psychiatric symptom severity—occurs completely via its relationship to the alliance.

This would also lend support to the idea that empathy and alliance are aspects of the same

therapeutic process. However, as discussed above, empathy and alliance were found to

be differentially correlated with other process variables.

Furthermore, indirect evidence from a recent study suggests that empathy and

alliance may be associated with somewhat different outcomes. Bedics, Henry, and

Atkins (2005) attempted to predict specific interpersonal outcomes using 3 subscales of

the VPPS that are broader than those used in current study. Two of subscales, referred to

as "Therapist Warmth" and "Patient Involvement" are of particular interest here. The

former is a combination of the Therapist Warmth and Friendliness scale used in the

current study and the Therapist Negative Attitude scale. The Patient Involvement scale is

the combination of the Patient Participation scale used in the current study and the Patient

Hostility scale. Bedics et al. found that Therapist Warmth at session 3 predicted an

increase in clients' self-reported affiliative behavior and a decrease in hostility toward a

significant other at termination. Patient Involvement at session 3 was not a significant

predictor of these interpersonal outcomes. Therapist Warmth (but not Patient

Involvement) at session 16 predicted decreases in clients' submissiveness with a

significant other. Patient Involvement at this session (but not Therapist Warmth)

predicted a decrease in clients' hostility toward a significant other. While not reported as

part of the current study, Therapist Empathy was found to be significantly correlated with

Therapist Warmth—but not with Patient Involvement—at session 3, $r = 0.49$, $p < 0.001$.

In contrast, Alliance was found to be significantly correlated with Patient Involvement—

but not with Therapist Warmth—both at session 3 ($r = 0.57$, $p < 0.01$) and at session 16 ($r = 0.51$, $p < 0.05$). This is important because it suggests not only that therapist empathy

and alliance are separable variables, but they may also be related to different types of

outcomes.

　　　　Another limitation of the current study stems from the fact that therapist empathy,

alliance, and other therapy process variables (i.e., VPPS scales) were all rated by

independent observers. This creates the possibility that correlations among these

variables are inflated due to common method variance. It should be noted, however, that

it is the pattern more than the magnitude of the relationships among these variables and

the outcome variable that is most relevant to the hypotheses being tested in this study. Furthermore, in the one exception to this—the magnitude of the correlation between empathy and alliance—any inflation of the relationship serves only as a more conservative test of the hypothesis that empathy and alliance are separable constructs.

Given the lack of a true control group in the current (as well as the original) study, it is not possible to determine the extent to which any decrease in symptoms was due to regression to the mean. Furthermore, given the lack of mid-treatment outcome measurement at session 3, the session when all major process variables in the current study were measured, it is not possible to rule out the possibility that early improvement or decline influenced the process variables at session 3, possibly accounting for any relationship between these variables and later outcome. However, some data are available that can provide indirect evidence against this possibility. First, while GSI scores at intake are not correlated with GAR scores at session 3, the latter are significant predictors of both termination GSI scores and residualized GSI scores. Furthermore, while Patient Psychic Distress—an observer-rated scale intended to tap general client symptomatology manifested in the treatment session—at session 3 is significantly correlated with symptoms (GSI) at both intake and termination, they (1) are not correlated with GAR scores and (2) controlling for session 3 Patient Psychic Distress scores does not diminish the relationship between GAR scores and residualized GSI scores. While not definitive, this at least indirect evidence that the relationship between alliance and outcome in the current study is not accounted for by early improvement in therapy. This is consistent with a recent study by Baldwin, Wampold & Imel (2007)

which disentangled the influence of early outcome, therapist effects, and client effects on the alliance-outcome relationship in a large sample. They concluded that the alliance-outcome relationship was largely due to the influence of the therapist, the other influences (early outcome and client effects) contributing little or nothing to the alliance-outcome relationship.

As noted by Horvath (1994) "the study of the relationship must ultimately focus on the specification of actions that promote and maintain the alliance to enable the therapist to use this knowledge for the benefit of the client" (p. 278). While the results of this study must be considered preliminary pending replication, they suggest some potentially important, if tentative, implications for clinical practice and training. The finding that observable therapist empathy behaviors were related to the alliance, which in turn predicted decreases in symptom severity at termination, points to specifiable behaviors, techniques, and/or attitudes that may be important in the formation of therapeutic alliances, which then presumably influence outcomes. The specifiability of these empathy behaviors makes it more likely that they can form the basis of clinical training and/or selection of effective therapists. Indeed, some have begun using the MEE in empathy training (Watson, personal communication, 12/28/06).

As the fertile field of therapeutic relationship research has turned its attention from broad relationships among broadly defined constructs to ever more precise specifications of more complex interpersonal processes, it has become more important than ever to carefully deconstruct our concepts, allowing their moving parts to move freely at nature's joints. It is becoming clearer that the alliance is such a concept in need

of deconstructing, freeing therapist and client contributions to the alliance from the central alliance concept itself. Failing to treat these concepts as separate processes, such as by combining them into single broad measures, serves to hinder further understanding of how these processes weave their respective ways through therapy toward success.

REFERENCES

Baldwin, S. A., Wampold, Bruce 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 & Clinical Psychology; 75 (6)*,
842-852.

Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in
social psychological research: Conceptual, strategic and statistical considerations.
Journal of Personality and Social Psychology, 51, 1173-1182.

Barrett-Lennard. G. T. (1962). Dimensions of therapist response as causal factors in
therapeutic change. *Psychological Monographs, 76(43)* (Whole No. 562.)

Bedics, J. D., Henry, W. P., and Atkins, D. C. (2005). The therapeutic process as a
predictor of change in patients' important relationships during time-limited
dynamic psychotherapy. *Psychotherapy: Theory, Research, Practice, Training, 42
(3)*, 279–284.

Beutler, L. E., Crago, M., & Arizmendi, T. G. (1986). Research on therapist variables in
psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of
psychotherapy and behavior change (3rd ed.*, pp. 257-310). New York: Wiley.

Bohart, A. C., & Greenberg, L. S. (1997). *Empathy reconsidered: New directions in
psychotherapy*. American Psychological Association, Washington, DC, US.

Bohart, A. C., & Greenberg, L. S. (1997). Empathy and psychotherapy: An introductory overview. In Bohart, A. C., & Greenberg, L. S. (1997). *Empathy reconsidered: New directions in psychotherapy*. American Psychological Association, Washington, DC, US.

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) Oxford University Press, London.

Bordin, E. S. (1979). The generalizability of the psychoanalytic concept of the working alliance. *Psychotherapy: Theory, Research & Practice, 16(3)*, 252-260.

Burns, D. D. (1989). *The feeling good handbook.* New York: William Morrow.

Burns, D. D. & Auerbach, A. (1996). Therapeutic empathy in cognitive-behavioral therapy: Does it really make a difference. In P. M. Salkovskis (Ed.). *Frontiers of cognitive therapy* (pp. 135-164). New York: Guilford Press.

Collins, L. M., Graham, J. W., & Flaherty, B. P. (1998). An alternative framework for defining mediation. *Multivariate Behavioral Research, 33*, 295-312.

Derogatis, L. R. (1983). *The SCLR-90: Administration, scoring, procedures manual-II*. Towson, MD: Clinical Psychometrics Research.

Duan, C. & Hill, C. E. (1996). The current state of empathy research. *Journal of Counseling Psychology, 43(3)*, 261-274.

Efron, B., & Tibshirani, R. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall/CRC.

Elliott, R., Watson, J. C., Goldman, R. N. & Greenberg, L. S. (2004). *Learning emotion-focused therapy: The process-experiential approach to change*. American Psychological Association, Washington, DC, US.

Freud, S. (1913). *On beginning the treatment: Further recommendations on the technique of psychoanalysis*. S.E., 12, 121-144.

Gaston, L., Marmar, C. R., Gallagher, D., Thompson, L. W. (1991). Alliance prediction of outcome beyond in-treatment symptomatic change as psychotherapy processes. *Psychotherapy Research, 1(2)*, 104-112.

Greenberg, L. S. & Pinsof, W. M. (1986). *The psychotherapeutic process: A research handbook*. Guilford Press, New York, NY.

Greenson, R. R. (1960). Empathy and its vicissitudes. *International Journal of Psychoanalysis, 41*, 418-424.

Guadagnoli, E., & Velicer, W. F. (1988). Relation of sample size to the stability of component patterns. *Psychological Bulletin, 103*, 265-275.

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' responses to training. *Journal of Consulting and Clinical Psychology, 61*, 441-447.

Henry, W. P., Strupp, H. H., Butler, S. F., Schacht, T. E., & Binder, J. L. (1993). Effects of training in time-limited psychotherapy: Changes in therapist behavior. *Journal of Consulting and Clinical Psychology, 61*, 434-440.

Horvath, A. O. (1981). An exploratory study of the working alliance its measurement and relationship to therapy outcome. *Dissertation*.

Horvath, A. O. (1994). Research on the alliance. In A. O. Horvath, and L. S. Greenberg (Eds.), *The working alliance: Theory, research, and practice* (pp. 259-286). John Wiley & Sons, Oxford, England.

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. 37-69). Oxford University Press, London.

Horvath, A. O. & Greenberg, L. S. (1986). *The development of the Working Alliance Inventory*. Guilford Press, New York, NY.

Horvath, A. O. & Greenberg, L. S. (1989). Development and validation of the Working Alliance Inventory. *Journal of Counseling Psychology, 36(2)*, 223-233.

Horvath, A. O. & Greenberg, L. S., (Eds.). (1994). *The working alliance: Theory, research, and practice*. John Wiley & Sons, Oxford, England.

Kazdin, A. E. (2007). Mediators and mechanisms of change in psychotherapy research. *Annual Review of Clinical Psychology, 3*, 1-27.

Linehan, M. M. (1993). *Cognitive-behavioral treatment of borderline personality disorder*. New York: Guilford Press.

MacKinnon, D. P., Krull, J. L., & Lockwood, C. M. (2000). Equivalence of the mediation, confounding, and suppression effect. *Prevention Science, 1*, 173-181.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G., & Sheets, V. (2002). A comparison of methods to test the significance of the mediated effect. *Psychological Methods, 7*, 83-104.

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(3)*: 438-450.

McCabe, R., & Priebe, S. (2004). The therapeutic relationship in the treatment of severe mental illness: A review of methods and findings. *The International Journal of Social Psychiatry, 50(2)*, 115-128.

Moseley, D. (1983). The therapeutic relationship and its association with outcome. *Unpublished master's thesis*, University of British Columbia, Vancouver, Canada.

O'Malley, S. S., Suh, C. S. & Strupp, H. H. (1983). The Vanderbilt Psychotherapy Process Scale: A report of the scale development and a process-outcome study. *Journal of Consulting and Clinical Psychology, 51*, 581-586.

Osipow, S. H., Carney, C. G., & Barak, A. (1976). A scale of educational-vocational undecidedness: A typological approach. *Journal of Vocational Behavior, 9(2)*, 233-243.

Raue, P.J. & Goldfried, M. R. (1994). *The therapeutic alliance in cognitive-behavior therapy*. John Wiley & Sons, Oxford, England.

Rogers, C. R. (1957). The necessary and sufficient conditions of therapeutic personality change. *Journal of Consulting Psychology, 21*, 95-103.

Rogers, C. R. (1975). Empathic: An unappreciated way of being. *Counseling Psychologist, 5(2)*, 2-10.

Safran, J. D. & Segal, Z. V. (1990). *Interpersonal process in cognitive therapy*. Jason Aronson, Inc, Lanham, MD.

Salvio, M., Beutler, L. E., Wood, J. M. & Engle, D. (1992). The strength of the therapeutic alliance in three treatments for depression. *Psychotherapy Research, 2(1)*, 31-36.

Shrout, P. E., & Bolger, N. (2002). Mediation in experimental and nonexperimental studies: New procedures and recommendations. *Psychological Methods, 7(4)*, 422-445.

Sobel, M. E. (1982). Asymptotic confidence intervals for indirect effects in structural equation models. In S. Leinhardt (Ed.), *Sociological Methodology*. (pp. 290-312). Washington DC: American Sociological Association.

Spielberger, C. D., Gorsuch, R. C., & Lushene, R. E. (1970). *Manual for the State Trait Anxiety Inventory*. Palo Alto, CA: Consulting Psychologists Press.

Steckley, P.L. (2006). An examination of the relationship between clients' attachment experiences, their internal working models of self and others, and therapists' empathy in the outcome of process-experiential and cognitive-behavioural therapies (Doctoral dissertation). Available from Dissertations and Theses database. (UMI No. AAT NR15730)

Strupp, H. H. (1993). The Vanderbilt Psychotherapy Studies: Synopsis. *Journal of Consulting and Clinical Psychology, 61(3)*, 431-433.

Strupp, H. H., Wallach, M. S. & Wogan, M. (1964). Psychotherapy experience in retrospect: Questionnaire survey of former patients and their therapists. *Psychological Monographs: General & Applied, 78(11*, Whole No. 588).

Tichenor, V. & Hill, C. E. (1989). A comparison of six measures of working alliance. *Psychotherapy, 26*, 195-199.

Truax, C. B., & Carkhuff, R. R. (1967). *Toward effective counseling and psychotherapy: Training and practice*. Aldine Publishing Co, Hawthorne, NY, US.

Waddington, L. (2002). The therapy relationship in cognitive therapy: A review. *Behavioural and Cognitive Psychotherapy, 30(02)*, 179-191.

Wang, V. & Anderson, T. Unpublished study of a manual for observer ratings of the Working Alliance Inventory.

Watson, J. C. (1999). Measure of expressed empathy. Unpublished manuscript. Department of Adult Education, Community Development, and Counseling Psychology, OISE, University of Toronto, Toronto, Ontario, Canada.

Watson, J. C. (2002). Re-visioning empathy. In D. J. Cain (Ed.), *Humanistic psychotherapies: Handbook of research and practice* (pp. 445-471). American Psychological Association, Washington, DC.

Watson, J. C. & Geller, S. M. (2005). The relation among the relationship conditions, working alliance, and outcome in both process-experiential and cognitive-behavioral psychotherapy. *Psychotherapy Research, 15(1-2)*, 25-33.

Watson, J. C. & Greenberg, L. S. (1994). The alliance in experiential therapy: Enacting the relationship conditions. In A. O. Horvath and L. S. Greenberg (Eds.), *The working alliance: Theory, research, and practice*. (pp. 153-172) John Wiley & Sons, Oxford, England.

Watson, J. C. & Greenberg, L. S. (1998). The therapeutic alliance in short-term humanistic and experiential therapies. In J. D. Safran and J. C. Muran (Eds.), *The therapeutic alliance in brief psychotherapy* (pp. 123-145). American Psychological Association, Washington, DC.

Watson, J. C. & Prosser, M. (2002). Development of an observer-rated measure of therapist empathy. In J. C. Watson, R. N. Goldman, and M. S. Warner (Eds.), *Client-centred and experiential psychotherapy in the 21st century* (pp. 303-314). Ross-on-Wye, UK: PCCS Books.

Watson, J. C., Prosser, M., Steckley, P., and Hiebert, D. (2003). Measure of Expressed Empathy Training Manual. Unpublished Manual, Department of Adult Education & Counseling Psychology, OISE/University of Toronto, Ontario, Canada.

Zetzel, E. (1956). Current concepts of transference. *International Journal of Psychoanalysis, 37*, 369-376.

Zuroff, D. C., Blatt, S. J., Sotsky, S. M., Krupnick, J. L., Martin, D. J., & Sanislow, C. A., et al. (2000). Relation of therapeutic alliance and perfectionism to outcome in brief outpatient treatment of depression. *Journal of Consulting & Clinical Psychology, 68(1)*, 114-124.

APPENDIX A: SAMPLE CHARACTERISTICS OF THE VANDERBILT II STUDY

Clients in the Vanderbilt II study were recruited via newspaper advertisements announcing low-cost therapy. Clients were selected who were judged to have clinically significant problems with a clear interpersonal component of severity comparable to ordinary outpatient samples (i.e., SCL-90-R Global Severity Index within one standard deviation of outpatient normative mean; Derogatis, 1983) and which did not indicate the need for other forms of treatment (e.g., substance use disorders, severe medical problems, or problems indicating the need for psychiatric medication or inpatient treatment). Eighty-four clients were finally accepted into the study, representing 19% of the potential client pool. Clients were considered dropouts if they completed less than 5 sessions. Out of the 84 clients who began the study, 4 (5%) dropped out and are not included in subsequent sample statistics or statistical analyses. The final sample was 77% female and 95% white, ranging in age from 24 to 64 (M = 41; SD = 10.4 years) with a mean education level of 3 years of college (range = 7th grade to doctorate). The mean number of sessions completed by the 80 included clients was 21.4 (SD = 6.1; range = 5 - 25). Clients were assigned diagnoses based on the Diagnostic and Statistical Manual of Psychological Disorders, third edition, by a trained interviewing clinician using the computerized version of the NIMH Diagnostic Interview Schedule. Eighty-seven percent of clients received an Axis I diagnosis and 67% received an Axis II diagnosis, with all patients qualifying for at least one Axis I or Axis II diagnosis. The mean Global Severity Index T-score (outpatient norms) from the SCL-90-R (Derogatis, 1983) was 48.1 (SD = 5.8). Sixty-eight percent of clients had participated in previous psychotherapy. Of the

original sample of 80, 16 of these were training cases and were dropped from subsequent analyses. Therefore, the final total study sample consisted of 64 clients.

APPENDIX B: RATER TRAINING AND RATING PROCEDURE FOR THE

MEASURE OF EXPRESSED EMPATHY AND THE TRUAX-CARKHUFF

ACCURATE EMPATHY SCALE

Training

Measure of Expressed Empathy. Raters trained to code videotape segments of

therapy sessions using an unpublished training manual developed by Watson, Prosser,

Steckley, and Hiebert (2003). Watson et al. provided "expert" ratings for each MEE item

for segments of published therapy training videotapes of well-known psychologists

demonstrating various types of individual psychotherapy. These ratings were used

initially in discussions of the MEE items individually as the raters attempted to

understand the meaning of each item as well as how different levels of each construct

being tapped manifested itself in therapy. Training then progressed to the point where

the 2 raters rated new segments of video separately and compared their ratings to the

expert ratings. Discrepancies were then discussed, at times with reference to literature

recommended by the main author of the measure (Watson, personal communication).

Once raters became comfortable, confident, and proficient at rating individual items, they

began to rate multiple segments using all items of the scale. These ratings and their sums

were then once again compared to the expert ratings and sums thereof. Discrepancies

were then discussed until a common understanding developed that was consistent with

the expert ratings. This process continued until high interrater reliability ($r > .85$) was

obtained several weeks in a row.

Truax-Carkhuff Accurate Empathy Scale. Since most studies using the AES were able to obtain reliable and valid AES ratings with minimally-trained raters, no training manual was used in this study for the training of raters on the AES. Instead, training piggybacked on the training for the MEE as described above, with the focus being on interrater agreement rather than matching some expert ratings (which were not available). Unfortunately, interrater reliability for the AES never reached a high level, either during training or after. However, the high correlation between the AES and the MEE provides some evidence for the validity of the scale.

Rating Procedure

All available videotapes of session 3 from the Vanderbilt II Psychotherapy Project video archive housed at Ohio University were coded using the MEE and the AES. The primary rater—the study author—rated 100% of the total number of videos, while the secondary rater—an advanced clinical psychology doctoral candidate—rated a stratified random sample of 50% of these same video segments. Ratings were done in batches of approximately 10 videotapes per week. Each week a list of the videotapes to be rated by the primary rater was made available to the second rater. From this batch, 50% of the videos were selected using a random number generator in Microsoft Excel. The primary rater was blind to which videos were being rated by the second rater. At the end of each week, the raters met to compare ratings and discuss discrepancies in order to prevent coder drift. None of the ratings were discarded, repeated, or combined. The final ratings used in the study were those of the primary coder. The ratings of the secondary coder were used only in the inter-rater reliability analyses.

APPENDIX C: TESTS OF REGRESSION ASSUMPTIONS FOR THE MEDIATION

MODEL

Before conducting the multiple-regression-based mediation analysis with residualized GSI scores as the criterion variable and MEE and GAR scores as predictors, several key statistical assumptions were tested for the model. The multiple regression analysis assumptions tested were (1) lack of influential multivariate outliers; (2) normality of the error distribution; (3) linearity of the relationship between criterion and predictor variables; and (4) independence of errors (no serial correlation).

Assumption: Lack of influential multivariate outliers
Tests: Cook's Distance; Mahalanobis Distance
Results:
- Cook's Distance: Highest value = 0.179 [critical value = 1]
- Mahalanobis Distance: Highest value = 7.143 [critical value (df = 3; $p < 0.001$) = 16.27]

Conclusion: Neither the highest Cook's Distance value nor the highest Mahalanobis Distance value exceeds its respective critical value. Therefore, there do not appear to be multivariate outliers in this data with respect to this regression model.

Assumption: Normality Distributed Error
Tests: Histogram of Residuals (from regression analysis); Normal Probability Plot of Standardized Residuals (from regression analysis)
Results: The following Histogram of Residuals shows the distribution of the residuals from the regression analysis. It appears to be roughly symmetrical and bell-shaped.

Histogram of Residuals
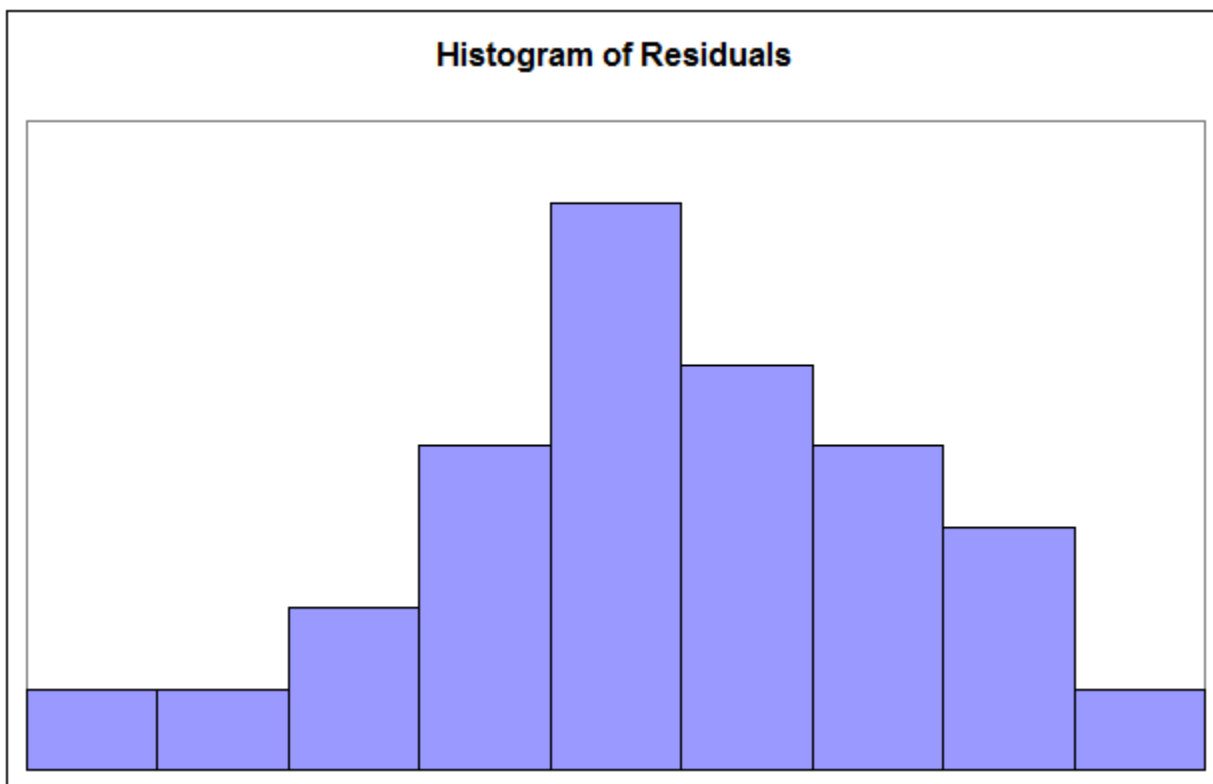
The following figure is the Normal Probability Plot of Regression Standardized Residuals. The plotted residuals fall largely close to the expected value line, though there appears to be a slight s-shape pattern, indicating possible kurtosis.

Normal P-P Plot of Regression Standardized Residual

Dependent Variable: GSI Residualized (z) Standardized Residual
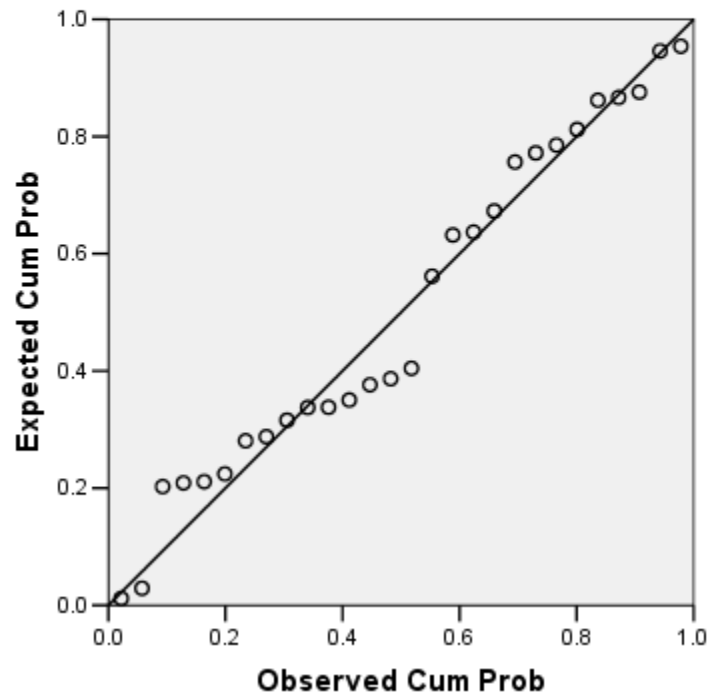
Conclusion: The assumption of the normality of errors appears to be sufficiently met. There is some slight indication of kurtosis, however.
Assumption: Linearity
Test: Scatterplot of Predicted Values vs. Residuals from Regression model

Scatterplot

Dependent Variable: GSI Residualized (z)   Standardized Residual
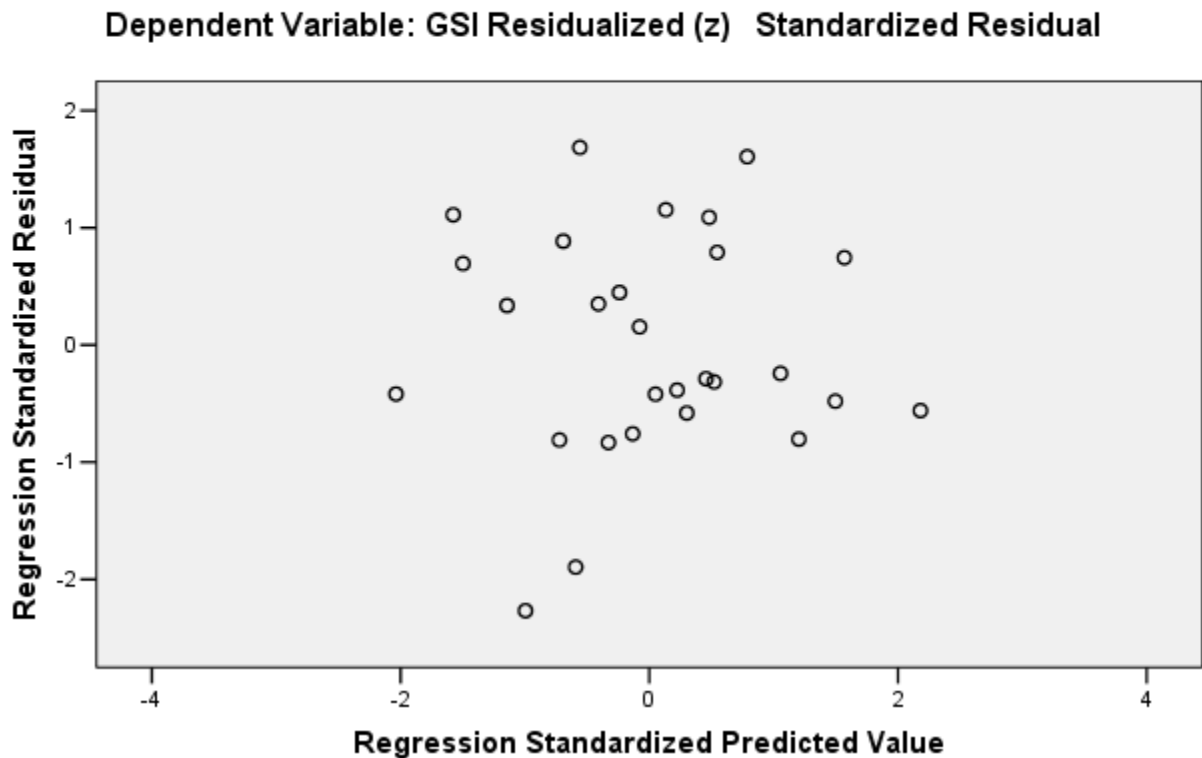
Results: The scatterplot below shows a roughly linear pattern with only minor apparent bowing. A test of fit of a quadratic relationship between the Predicted Values and Residuals showed a very small effect size ($R^2 = 0.002$). Conclusion: The relationship between predictors and criterion variable in this regression model is largely linear with an only very minor curvilinear relationship. Therefore, the assumption of linearity is largely met.

Assumption: Independence
Test: Durbin-Watson test of autocorrelation
Results: Durbin-Watson = 1.639 (critical value = 1.65)
Conclusion: The observed value of the Durbin-Watson statistic does not exceed the critical value for this model and data set, suggesting that the assumption of independence of errors (no serial correlation) has not been violated.

APPENDIX D: CORRELATIONS AMONG MAJOR STUDY VARIABLES, ALTERNATIVE EMPATHY AND ALLIANCE

SCALES, AND VPPS SCALES.

	N	AES	WAI	GAR	THWF	NETA	THEX	PTPA	PTEX	PTPD	PTHO	PTDP
MEE	54	0.75***	0.41*	0.55**	0.52**	0.03	0.08	0.2	-0.09	-0.27*	-0.35*	-0.26
AES	54		0.42*	0.55**	0.33**	0.19	0.21	0.08	0.07	-0.11	-0.28*	-0.08
WAI	30			0.87***	0.3	-0.02	0.04	0.54**	0.28	0.17	-0.55**	-0.13
GAR	30				0.13	0.17	-0.07	0.5**	0.26	0.12	-0.45*	-0.23
THWF	54					0.08	0.42**	0.15	0.04	-0.11	-0.33*	-0.03
NETA	54						0.17	-0.05	0.2	-0.05	0.07	0.08
THEX	54							-0.18	0.37**	0.2	0.12	0.42**
PTPA	54								0.54***	0.14	-0.25	-0.15
PTEX	54									0.50***	-0.05	0.35*
PTPD	54										-0.23	0.22
PTHO	54											0.23

Note: MEE = Measure of Expressed Empathy; AES = Truax-Carkhuff Accurate Empathy Scale; WAI = Working Alliance Inventory (manualized observer rating); GAR = Global Alliance Rating; THWF = VPPS Therapist Warmth & Friendliness scale; NETA = VPPS Negative Therapist Attitude scale; THEX = VPPS Therapist Exploration scale; PTPA = VPPS Patient Participation scale; PTEX = VPPS Patient Exploration scale; PTPD = VPPS Patient Psychic Distress scale; PTHO = VPPS Patient Hostility scale; PTDP = VPPS Patient Dependence scale; sample sizes are N = 54 except correlations with GAR or WAI, the latter having samples sizes of N = 30; *p < .05; **p < .01; ***p < .001

Table 1. Correlations among Psychotherapy Process Variables.

	MEE	GAR	THWF	PTPA
MEE	-	0.55** (30)	0.52** (54)	0.2 (54)
GAR		-	0.13 (30)	0.5** (30)
THWF			-	0.15 (54)
PTPA				-

Note: MEE = Measure of Expressed Empathy; GAR = Global Alliance Rating; THWF = VPPS Therapist Warmth & Friendliness scale; PTPA = VPPS Patient Participation scale; respective sample sizes in parentheses; *p < .05; **p < .01; ***p < .001

Table 2. Sample Sizes, Means and Standard Deviations of Study Variables.

Variable	N	Mean	SD
Measure of Expressed Empathy (MEE)	54	5.08	0.84
Accurate Empathy Scale (AES)	54	2.46	1.19
Working Alliance Inventory-manualized observer rating (WAI-OM)	30	606.83	19.17
Global Alliance Rating (GAR)	30	19.82	3.59
VPPS Patient Participation (PTPA)	54	322.04	32.05
VPPS Therapist Warmth & Friendliness (THWF)	54	246.16	54.88
Global Severity Index (GSI) residualized change score	52	0	1

Table 3. Bootstrap Analyses of the Magnitude and Statistical Significance of the Direct and Indirect Effects (MEE & GAR).

Effect	B	SE	t	p
Effect of Empathy (MEE) on Alliance (GAR)	2.71**	0.81	3.35	0.0025
Direct Effect of Alliance on Symptom Change	-0.17**	0.05	-3.18	0.0039
Total Effect of Empathy on Symptom Change	-0.01	0.25	-0.05	0.9624
Direct Effect of Empathy on Symptom Change	0.44	0.26	1.70	0.1012
	B	SE	95% CI	
Indirect Effect of Empathy on Symptom Change (through Alliance)	-2.68*	1.47	-6.11, -0.46	

NOTE: N = 28; *$p < 0.05$; **$p < 0.01$; negative Symptom Change scores mean reduction in symptoms from pre-treatment to termination.

Table 4. Bootstrap Analyses of the Magnitude and Statistical Significance of the Direct and Indirect Effects (AES & GAR).

Effect	B	SE	t	p
Effect of Empathy (AES) on Alliance (GAR)	2.23**	0.66	3.39	0.0022
Direct Effect of Alliance on Symptom Change	-0.16**	0.05	-2.95	0.0068
Total Effect of Empathy on Symptom Change	-0.06	0.2	-0.29	0.7775
Direct Effect of Empathy on Symptom Change	0.29	0.22	1.36	0.1849
	B	SE		95% CI
Indirect Effect of Empathy on Symptom Change (through alliance)	-0.36*	0.15		-0.76, -0.14

NOTE: N = 28; *$p < 0.05$; **$p < 0.01$; negative Symptom Change scores mean reduction in symptoms from pre-treatment to termination.

Table 5. Bootstrap Analyses of the Magnitude and Statistical Significance of the Direct and Indirect Effects (MEE & WAI-OM).

Effect	B	SE	t	p
Effect of Empathy (MEE) on Alliance (WAI-OM)	10.79*	4.63	2.33	0.0279
Direct Effect of Alliance on Symptom Change	-0.03*	0.01	-2.62	0.0148
Total Effect of Empathy on Symptom Change	-0.12	0.25	-0.05	0.9624
Direct Effect of Empathy on Symptom Change	0.26	0.25	1.04	0.3083
	B	SE	95% CI	
Indirect Effect of Empathy on Symptom Change (through alliance)	-0.30*	0.19	-0.74, -0.02	

NOTE: N = 28; *$p < 0.05$; **$p < 0.01$; negative Symptom Change scores mean reduction in symptoms from pre-treatment to termination.

Table 6. Bootstrap Analyses of the Magnitude and Statistical Significance of the Direct and Indirect Effects (AES & WAI-OM).

Effect	B	SE	t	p
Effect of Empathy (AES) on Alliance (WAI-OM)	9.14*	3.77	2.43	0.0225
Direct Effect of Alliance on Symptom Change	-0.02*	0.01	-2.49	0.0198
Total Effect of Empathy on Symptom Change	-0.06	0.20	-0.29	0.7775
Direct Effect of Empathy on Symptom Change	0.16	0.21	0.79	0.4385
	B	SE	95% CI	
Indirect Effect of Empathy on Symptom Change (through alliance)	-0.23*	0.14	-0.58, -0.03	

NOTE: N = 28; *p < 0.05; **p < 0.01; negative Symptom Change scores mean reduction in symptoms from pre-treatment to termination.