1977

The relationship between therapist empathy in therapy and nontherapy settings and some contributing components to empathic understanding

Ted William Grubb
University of Massachusetts Amherst

Follow this and additional works at: https://scholarworks.umass.edu/theses

Retrieved from https://scholarworks.umass.edu/theses/1575

This thesis is brought to you for free and open access by ScholarWorks@UMass Amherst. It has been accepted for inclusion in Masters Theses 1911 - February 2014 by an authorized administrator of ScholarWorks@UMass Amherst. For more information, please contact scholarworks@library.umass.edu.
THE RELATIONSHIP BETWEEN THERAPIST EMPATHY IN THERAPY AND NONTHERAPY SETTINGS AND SOME CONTRIBUTING COMPONENTS TO EMPATHIC UNDERSTANDING

A Thesis Presented
By
TED WILLIAM GRUBB

Submitted to the Graduate School of the University of Massachusetts in partial fulfillment of the requirements for the degree of MASTER OF SCIENCE
September 1977
Department of Psychology
THE RELATIONSHIP BETWEEN THERAPIST EMPATHY
IN THERAPY AND NONTHERAPY SETTINGS AND SOME
CONTRIBUTING COMPONENTS TO EMPATHIC UNDERSTANDING

A Thesis Presented
By
TED WILLIAM GRUBB

Approved as to style and content by:

Richard F. Haase, Chairperson of Committee

Melinda A. Novak
Melinda Novak, Member

Bonnie Strickland
Bonnie Strickland, Member

Castellano Turner
Castellano Turner, Member

Bonnie Strickland, Department Head
Psychology Department
Acknowledgements

I am deeply grateful to the committee members, Melinda Novak, Bonnie Strickland and Cass Turner, for their time, commitment and encouragement throughout all phases of this research project.

I would also like to thank Norman Watt for his hours of patient and meticulous attention spent in helping to create the basic design of the study.

Dick Haase was invaluable for both his personal support and insightful comments on the treatment and organization of the data. His thoughtful views allowed me to maintain a sound perspective despite occasional frustrations in the execution of the study.

In retrospect, I believe that the most satisfaction for me has come from the transformation of my views toward scientific research and the pursuit of knowledge. From a condition of fear and ignorance, I have moved to a personal sense of confidence and competency in what I do. It has not been a matter of arriving at a final destination, but rather of discovering a process in which I feel comfortable and can anticipate further growth. I am indebted to all of the above-mentioned people, who helped to facilitate this discovery.

I would also like to thank the therapists and clients in the Psychological Services Center who graciously cooperated in this study.

Finally, I wish to make special mention of my parents. They have consistently provided a gentle mixture of prodding, nurturance, and acceptance and they continue to be important influences in my personal and professional life.
# TABLE OF CONTENTS

Acknowledgements iii  
List of Tables vi  

## CHAPTER I  INTRODUCTION AND LITERATURE REVIEW  
Empathy: A Personality Characteristic? 3  
Problems In the Research 6  
Predictive Measures of Empathy 7  
Self-report Measures of Empathy 14  
Hogan’s Empathy Scale: a Paper-and-pencil Test 17  

## CHAPTER II  RATIONALE  
Empathy Components 20  
Major Hypotheses 22  

## CHAPTER III  METHOD  
Subjects 27  
Raters 27  
Measures 29  
Reliability of Measures 30  
Procedure 31  

## CHAPTER IV  RESULTS  
Reliability and Agreement Between Professional Associates 34  
Major Hypotheses 39  

## CHAPTER V  DISCUSSION  
Problems in Design: Major Hypotheses 47  
Additional Problems in Design 50  
Conclusions and Suggestions for Future Research 53
# TABLE OF CONTENTS
(Continued)

## REFERENCES

## APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Appendix A</td>
<td>Subject Consent Form</td>
<td>61</td>
</tr>
<tr>
<td>Appendix B</td>
<td>Rating Scales for Empathy Components</td>
<td>62</td>
</tr>
<tr>
<td>Appendix C</td>
<td>Hogan Scale of Empathy</td>
<td>70</td>
</tr>
<tr>
<td>Appendix D</td>
<td>Progress/Outcome Scale: Therapist Form</td>
<td>74</td>
</tr>
<tr>
<td>Appendix D</td>
<td>Progress/Outcome Scale: Client Form</td>
<td>75</td>
</tr>
</tbody>
</table>
List of Tables

Table 1, An Example of a Multitrait-multimethod matrix 24

Table 2, Multitrait-multimethod matrix: Professional Associates X Professional Associates 35

Table 3, Convergent and Discriminant Validity Between Professional Associates 37

Table 4, Average Reliabilities and Agreements Between Professional Associates 39

Table 5, Pearson r correlations of Therapist Composite Empathy Ratings by Professional Associates and by Clients 40

Table 6, Pearson r correlations of Therapist and Client Outcome with Therapists' Composite Empathy Ratings by Clients and Professional Associates 41

Table 7, Multitrait-multimethod Matrix: Professional Associates X Clients 43

Table 8, Rank-order Correlations of Empathy Components by Professional Associates and Clients 45
CHAPTER I
INTRODUCTION AND LITERATURE REVIEW

Since Rogers (1957) identified empathy as one of the "necessary and sufficient" conditions for change to occur in psychotherapy, a tremendous amount of writing and research has attempted to explain empathy. In clinical writings, there appears to be a fairly clear consensus on what we mean when we talk about empathy. Rogers states that "to sense the client's private world as if it were your own, but without ever losing the 'as if' quality--this is empathy" (1957, p. 99). At its most basic level, empathy is a process of understanding; most clinicians seem to agree that empathic understanding is a process of seeing the world through the clients' eyes, to perceive it as they do, and to communicate this understanding.

In spite of the apparent clinical consensus on the meaning of empathy, the research literature is plagued with problems. The difficulties center around the lack of construct validity in the tests used to measure empathy. Cronbach and Meehl (1955), in a discussion of construct validity, note that "unless substantially the same nomological net is accepted by the several users of the construct, public validation is impossible" (p. 291). In empathy research, there is a great deal of divergence on assumptions about the assessment of empathy; there is no universally accepted operational definition of empathy.

One major difficulty in empathy research is the question of whether an empathy measure reflects a general ability, like a per-
sonality trait, or a temporary condition, subject to a greater degree of variation. There is no agreement on this issue and the method used to measure empathy in each case reflects the belief held by the particular researcher. The lack of unity among researchers on this basic assumption about empathy confounds comparative analysis.

A second problem is the lack of specificity in most studies in their use of the construct of empathy. The empathy process is a rather complex, multifaceted one, and calling it simply "empathy" without any further specification does not help to clarify, but may indeed confuse the issue. Cronbach and Meehl emphasize that "'learning more about' a theoretical construct is a matter of elaborating the nomological network in which it occurs, or of increasing the definiteness of the components" (1955, p. 290). Although the empathy process may be said to be made up of many components, these are usually left unspecified operationally and are collectively referred to as "empathy".

This study examined the relationship between some empathy-related behaviors of therapists outside of therapy, and the level of empathic understanding they displayed in therapy. The validity of two types of personality measures was examined: 1. questionnaire (the Hogan Empathy Scale); 2. ratings of therapists made by professional associates based on knowledge of the therapists' behavior outside of therapy. Clients' ratings of empathy served as the validity criterion. Additionally, there was a correlational analysis of empathy and outcome as perceived by therapist and client.
Empathy: A Personality Characteristic?

In his earlier writings, Rogers (1957, 1959) states that empathy does not necessarily have to be characteristic of a therapist's life in general; the critical thing is for the therapist to be empathic during the hour of therapy with the client. In a recent paper, however, Rogers (1975) implies that empathy may reflect a more characterological trait, being affected by the psychological maturity of the therapist.

Truax and Carkhuff (1967), in an extensive number of studies on therapeutic process and outcome, conclude that the critical core of facilitative conditions offered by a therapist indicate a general personality characteristic of the therapist. Bergin and Strupp (1972) suggest that a therapist's basic personality disposition influences therapeutic interaction and patient outcome; these basic dispositions are seen as largely inseparable from any technique or specific mode of therapy.

The Empathy Scale developed by Hogan (1969) also assumes that empathy is a general characterological disposition. It is a 64-item, true-false questionnaire, and is said to assess a disposition to act in a moral way, to empathically see another's point of view.

There may be some confusion over the use of the term "personality characteristic". Personality is defined by Holt as a "patterned set of dispositions to behave in particular ways" (1969, p. 6). One possible inference (to the extreme) from this definition is that a person's behavior is consistent across various situations; a valid measure of his/her behavior in certain situations will allow
reliable prediction of his/her behavior in other situations. In reference to therapists' behavior, this would indicate that they are disposed to act in similar ways both in and outside of therapy.

An argument to the opposite extreme might also be made which still satisfies Holt's definition of personality: a person's behavior is consistent but only within similar situations; a measure of his/her behavior in one situation is predictive only for other highly similar situations. For therapists then, their behavior in therapy might not be indicative of their behavior outside of therapy.

A more accurate description of personality probably lies somewhere between the above two extremes. A certain amount of consistency in behavior may be expected across situations, although varying external conditions may cause or allow certain changes in behavioral patterns. For therapists, a disposition to act in certain ways may be identified in their behavior both in and out of therapy. The use of the term "personality characteristic" may be confusing if it is not specified whether the behavior referred to is inside or outside the therapy session. An important question here is: to what degree are therapists' behavior patterns outside of therapy related to their behavior patterns in therapy?

It would follow from the conclusions of Rogers, Truax and Carkhuff that there is indeed a relationship between therapists' behavior in and out of therapy. An alternate position might hold that empathic ability in therapy should not necessarily be affected by the way a therapist behaves in general. The empathy process might be a temporary condition, a therapy technique which can be
effectively used in specific situations (therapy). The studies by Rogers, Truax and Carkhuff imply that empathy is a general behavioral characteristic of therapists. They do not, however, differentiate the therapy room from the multitude of settings in which therapists' personalities may be measured. Thus, instead of asking whether or not empathy is a personality characteristic, it is perhaps better to examine how therapists' empathy-related behaviors outside of therapy correlate with their empathy-related behaviors in therapy. An initial attempt at this examination was made in this study.

The distinction between therapy and non-therapy behavior is not usually made by researchers when making extrapolations from research data to describing empathic therapists. Truax and Carkhuff, for example, infer from data collected in therapy or pseudo-therapy settings that their descriptions of therapists also apply to therapists' general (non-therapy) attitudes and behavior. Hogan assumes that therapists' self-reported descriptions of their empathic behavior outside of therapy will be related to their behavior in therapy. The empirical evidence to support these assumptions has not been presented.

Thus, the most commonly used empathy measures have failed to adequately examine the relationship between therapists' in-therapy behavior with their behavior outside of therapy. Additionally, these measures still suffer from their lack of construct validity which will be more specifically discussed later.
Problems In the Research

Many of the studies on empathy published to date are plagued by methodological problems. Design and statistical considerations limit the worth of some studies, while inconsistencies in the treatment or measurement of empathy from one study to the next (often in those by the same researcher) make comparisons frustrating and sometimes meaningless. Therefore, great care must be taken before making any generalizations from one set of statements to another regarding empathy; two authors may use the term "empathy" when they are actually referring to two very different things.

There are three basic ways in which empathy has been measured in psychotherapy: the therapist tries to predict the client's responses to a set of self-descriptive items; the therapist-client interaction is rated by trained judges; the therapist and client make a self-report of their experience. Each of these measures reflects the different assumptions about empathy made by the respective researchers and will be discussed later.

Another class of empathy measures have also been developed which do not rely upon measurement of a specific interaction between two people for an empathy rating or score. These are paper-and-pencil tests which identify empathy as a characterological trait, measurable by personality inventories which are similar to or consist in part of such instruments as the MMPI.

Certain of these measures have achieved some degree of popularity in psychotherapy research. Reliabilities of .50 to
90 have been reported on most of these and they have been shown to correlate positively with various measures of therapeutic outcome. There remains, however, the fact that no available measure of empathy has demonstrated full discriminant validity. Although these measures appear to measure something, there is no clear evidence that this something can validly be referred to as empathy.

In the following sections, each of the most common empathy measures will be briefly discussed. It will be noticed that the clinical definitions of empathy given by the authors of each measure are very similar. The working or operational definition of empathy in each case, however, is quite different. Specific methodological problems with each measure will also be noted.

Predictive Measures of Empathy

Dymond (1949, 1950) developed one of the first measures of empathy which relied upon some specific interaction between two or more people. Her definition of empathy gives some indication of the process her scale was designed to measure: "the imaginative transposing of oneself into the thinking, feeling, and acting of another and so structuring the world as he does" (Dymond, 1949, p. 127). The scale consisted of a five-point continuum for each of six characteristics: self-confidence, superior-inferior, selfish-unselfish, friendly-unfriendly, leader-follower, sense of humor. Following a brief interaction, two people would use the scale to rate the following: self, other, other as rated by other, self as rated by other. Empathic ability was seen as the
degree to which one person could accurately predict the ratings made by the other person.

The scale was subsequently used by Hastorf and Bender (1952; Bender & Hastorf, 1953) who found that it was useful in identifying people who could, with some consistency, make empathic predictions about others. They also noted, however, that these predictive scores of empathy were confounded by projection on the part of the predictor. The problem was that the prediction method did not tell whether the subject had achieved true empathic understanding, or was merely relying on self reference for their ratings. If the rater happened, by chance, to have similar personal experiences and perceptions to the person they were rating, then it would appear that a high degree of understanding existed. The discrepancy became evident when the supposedly high empathy person had to predict for persons of dissimilar background. It became clear that the predictions were not truly empathic, but were based on the person's projections or self reference.

Cronbach (1955) criticized predictive measures of empathy because they did not differentiate between true understanding and stereotypy or self reference. Halpern (1955) used a predictive measure of empathy and found a positive correlation between the empathy score and the similarity of predictor and predictee. Various attempts were made by Dymond and others to correct for projection and assumed similarity between raters, but a satisfactory answer to the problem was never found.

Another difficulty with Dymond's scale was that it was
limited to only six characteristics, which did not allow for fine discriminations of personal description. The development of the Q-sort technique (Block, 1961; Butler & Haigh, 1954) offered an instrument which could make the finer discriminations between personality characteristics, with 70 to 100 adjectives to be sorted. Rogers (1957) suggested that a therapist's match of their client's self Q-sort would be one way to measure the level of empathy achieved by the therapist.

Although the Q-sort technique did allow for a more complex matching of the client's cognitive structuring of the world, it still suffered from the problem of projection common to all predictive measures of empathy. In a therapeutic sense, the predictive measures of empathy had even more serious problems. With their emphasis on the cognitive structuring of the world by the client, predictive empathy measures largely ignored the emotional experience of both therapist and client in the relationship. The ability of the therapist to cognitively match the client's self-description was not necessarily indicative of anything emotional or therapeutic in the relationship; there was no attention to the ability of the therapist to effectively communicate his/her empathic understanding to the client.

In the same article in which he suggested matching Qsorts to measure empathy, Rogers (1957) actually defined empathy as a much more complex process than could ever be measured by a predictive instrument. In the clearest statement of his theory of psychotherapy to that time, Rogers listed empathy as one of the
"necessary and sufficient" conditions for constructive personality change in therapy. It is important here to note Rogers' definition, because it was his formulation of the empathy process which led to the subsequent development of two new ways to measure empathy: the rating of the therapeutic interaction by trained judges, and self reports of the therapist and client. Rogers stated that empathy was the ability "to sense the client's private world as if it were your own" (1957, p. 99), and further specified:

The final condition as stated is that the client perceives, to a minimal degree, the acceptance and empathy which the therapist experiences for him. Unless some communication of these attitudes has been achieved, then such attitudes do not exist in the relationship as far as the client is concerned, and the therapeutic process could not, by our hypothesis, be initiated (1957, p. 99).

In the initial description of empathy as the ability to accurately sense another person's private world, Rogers did not differ significantly from the definition by Dymond already mentioned. In his subsequent clarification, however, Rogers added two very important points to the basic condition of understanding: 1. the therapist must attempt to communicate this understanding to the client; 2. the client must perceive the empathic understanding offered by the therapist.

Trained Judges' Ratings of Therapist-offered Conditions

C. B. Truax, a student of Rogers', developed the Truax
Accurate Empathy (AE) Scale (Truax, 1961; see also Truax & Carkhuff, 1967). The scale consisted of a nine point continuum, from no empathy to high empathy, and was scored by trained judges who listened to random segments of tape-recorded interviews. The judges did not necessarily have to have any knowledge of psychology or psychotherapy (many of the judges used were naive undergraduates), but were trained specifically in the use of the scale.

In his introduction to the AE scale, Truax stated that "accurate empathy involves both the therapist's sensitivity to current feelings and his verbal facility to communicate this understanding in a language attuned to the client's current feelings" (Truax & Carkhuff, 1967, p. 46). It is clear that Truax meant for his AE scale to attend to the ability of the therapist to communicate in an empathic manner to the client. There was, however, no attention given to how the client perceived the message offered by the therapist. Judges were trained to attend to the therapist's responses, and to ignore as much as possible, the client's statements. In fact, one study by Truax (1966a) found that judges' ratings of tapes with client's responses edited out matched those of judges who rated unedited versions of the same tapes. Truax concluded that AE measurements were not "contaminated" by client responses.

The Truax AE scale has achieved some success as a predictor of success in therapy. Truax and Carkhuff (1967) reviewed 28 studies and found that high scores by a therapist on the AE scale was a reliable predictor of positive outcome in therapy. The
Wisconsin Schizophrenia Project, an extensive study of therapy and outcome with hospitalized schizophrenics, was completed by Rogers and his associates (Rogers et al., 1967); their findings corroborated those of Truax and Carkhuff regarding the AE scale. Truax (1970) and Mullen and Abeles (1971) also completed outcome studies which confirmed the predictive validity of the AE scale.

One problem with the AE scale is that judges must depend strictly upon the verbal communication of the therapist and ignore the non-verbal component. Shapiro, Foster, and Powell (1968) found that empathy can be accurately judged from strictly non-verbal cues, with facial gestures being especially important. Haase and Tepper (1972) found that lack of eye contact, backward trunk lean, and spatial distancing from a client can reduce high levels of therapist verbal empathy to perceived messages of non-empathy. It seems clear from these findings that the non-verbal behavior of the therapist is a critical part of his/her attempt to communicate empathic understanding.

Several criticisms of the AE scale were made by Chinsky and Rappaport (1970; rejoinder by Truax, 1972; reply by Rappaport and Chinsky, 1972). They noted that the reported reliability coefficients may be spuriously inflated by the continued ratings of therapists by the same judges, and that the ratings may depend largely on the type of training given to the judges. They also suggested that, in light of the raters' attention to only the therapists' responses, the raters may have been attending to the voice quality,
or some other characteristic of the therapists besides empathy. They concluded that the Truax AE scale lacks discriminant validity; it may reflect a more general therapist quality of interest and involvement in therapy.

The lack of discriminant validity in the Truax AE measure is evident in many of the studies which also use a similar rating scale for genuineness and nonposessive warmth (positive regard). These three scales often have very high correlations with each other (Truax & Carkhuff, 1967). Muehlberg et al. (1969) reported the inter-correlations between judges' ratings of empathy, regard, genuineness, concreteness, and self-disclosure: empathy ratings correlated from .78 to .91 with the other four ratings. Mintz et al. performed a factor analysis on judges' ratings of 110 therapist and client variables. They found that items on therapist perceptiveness, security, skill, empathy, acceptance, reassurance and warmth all loaded to the order of .75 to .86 on one factor: optimal empathic relationship. Empathy seems to be among several qualities which combine to describe the "goodness" of therapy, but judges' ratings of empathy have not been able to distinguish it from these other general qualities.

Another serious criticism of the Truax AE scale is that it fails to attend to the client's statements or perceptions of the therapy situation. The accuracy of any empathic statement by the therapist must be suspect in the absence of the client's response to the communicated message. It will be remembered that according to Rogers, the perception by the client of the
therapist's empathic understanding is the determinant of whether or not accurate empathy actually exists in the relationship. The importance of the client's perceptions of the relationship led to the development of the third type of empathy measure, the self-report questionnaire.

Self-report Measures of Empathy

Several measures have been developed which allow a client to rate the degree to which certain conditions have been offered by their therapist in therapy (Comrey, Backer, & Glaser, 1973). One of these measures, the Barrett-Lennard (1962) Relationship Inventory (RI), was developed from Rogers' conception of the conditions necessary for therapeutic change. It was designed to measure five dimensions of the therapeutic relationship: empathy, congruence, level of regard, unconditionality of regard, willingness to be known.

In adding further emphasis to Rogers' (1957) stated position, Barrett-Lennard wrote that "the client's experience of his therapist's response is the primary locus of therapeutic influence in their relationship . . . It is what the client himself experiences that affects him directly" (1962, p. 2). Barrett-Lennard acknowledged that the conscious perceptions of the client might not be a perfect indication of the client's experience, but concluded that the client's own report was "the most direct and reliable evidence" available of the actual experience.

The RI was intended as a more global measure of the therapeutic
relationship than was Truax' AE scale. The AE scale was usually applied to brief (2 to 4 min.) segments of initial or termination interviews. The RI was not administered until after the therapist and client had met for at least five sessions. This allowed the relationship to achieve a relative amount of stability, and the client responded to the general condition that he or she had received from the therapist, without reference to any particular interview or part of an interview.

The RI has reported high test-retest reliabilities of .74 to .92, and has been used in over 60 studies. The predictive validity of the RI has been shown in several studies (Barrett-Lennard, 1962; Hansen et al., 1968; Kurtz & Grummon, 1972; Rogers et al., 1967; Stoffer, 1968) which reported significant positive correlations between clients' RI ratings and therapeutic outcome as measured by a number of criteria.

There is disagreement in the literature as to whether or not clients' RI ratings correlate with judges' ratings on the Truax AE scale. Several studies (Caracena & Vicory, 1969; Fish, 1970; McWhirter, 1973; Truax, 1966b) found no correlation between clients' perceptions and judges' ratings of therapist-offered conditions. There are, however, questionable factors in some of these studies. The Caracena and Vicory (1969) and McWhirter (1973) studies used undergraduate volunteers in a pseudo-therapeutic setting, and the RI was filled out following only one or two brief (30 to 40 min.) meetings between the participants. (Barrett-Lennard had suggested it was necessary to have at least five sessions between a therapist
and client before the RI was administered). The Truax (1966b) study involved clients in groups, rather than individual therapy, which may have had a confounding effect on the clients' ratings of conditions they received from the therapist.

Truax and Carkhuff (1967) reported studies in which they compared AE ratings with Truax' Relationship Questionnaire, an instrument developed from and similar to the Barrett-Lennard RI. They concluded that perceptions of non-hospitalized clients were in agreement with judges AE ratings and predicted outcome. With hospitalized clients, however, they found that more severely disturbed patients were inaccurate in their perceptions of therapeutic conditions as measured by the AE scale. Truax and Carkhuff concluded that, generally, client perceptions of therapeutic conditions are unreliable, and that clients often do not know what is best for them. This conclusion is, of course, greatly at odds with Rogers' and Barrett-Lennard's contention that the client's experience is the reality which makes a difference in therapy.

In contrast with the Truax and Carkhuff (1967) conclusions regarding the incorrectness of clients' perceptions of therapy, Rogers et al. (1967) noted that "our unbiased raters and our schizophrenic patients tended to make similar evaluations of the therapeutic relationship" (p. 77). Bozarth and Grace (1970) studied 15 experienced counselors and their clients in university counseling centers throughout the nation. They found that the empathy subscale of the RI correlated .36 (ns) with the Truax AE scale. The total RI correlated .47 (p < .05) with the Truax AE scale.
No clear statement can be made as to whether or not clients' RI ratings agree with judges' ratings on the AE scale. There is some evidence for the validity of the RI in its ability to predict therapeutic outcome. Like the AE scale, however, the RI has not demonstrated discriminant validity.

Walker and Little (1969) conducted a factor analysis of the RI, and found low and nonsignificant correlations between the unconditionality and level of regard subscales, which supported the operational separation of those two variables. They found significant correlations between the empathy and congruence subscales, and concluded that these two scales were empirically indistinguishable. This supported the original scores obtained by Barrett-Lennard, with a correlation of .85 between the congruence and empathy subscales. Barrett-Lennard argued that a high correlation between the two should be expected, since congruence is a primary factor in a person's potential for empathic understanding. This, however, again demonstrates the lack of specificity in the empathy construct as measured by this instrument.

Hogan's Empathy Scale: a Paper-and-pencil Test

The Empathy Scale developed by Hogan (1969) was based on a definition of empathy similar to Dymond's (1949) and Rogers' (1957) definitions. Hogan defined empathy as "the intellectual or imaginative apprehension of another's condition or state of mind without actually experiencing that person's feelings" (1969, p. 308).

To construct the Empathy Scale, Hogan asked seven psychologists...
to describe a highly empathic person from the given definition, using the 100-item California Q-sort (Block, 1961); their composite description was the empathy criterion. He then had 8 to 10 skilled observers describe 211 subjects with the Q-sort following weekend live-in sessions; the correlation of the subjects' Q-sort with the empathy criterion was their empathy score. Subjects were then divided into high, medium, and low empathy groups and were tested on the California Personality Inventory (CPI), the MMPI, and a set of items from the University of California's Institute of Personality Assessment and Research (IPAR). Their scores were compared and a 64-item scale (32 true, 32 false) was developed, with 31 CPI items, 25 MMPI items, and 8 IPAR items.

Hogan actually developed the Empathy Scale from a non-psychotherapy perspective. He has described empathy as one aspect of a person's character structure pertaining to moral behavior; it is a general characterological trait, which allows the person to consider the implications of his/her behavior for others, and to adopt a moral point of view (Hogan, 1973; Hogan & Dickstein, 1972). Hogan described the empathic person as one who is extraverted, non-neurotic, socially acute, and sensitive to nuances in personal behavior (1969).

Several recent studies have utilized the Hogan Empathy Scale and have found that it correlates positively with other tests on field independence (Martin & Toomey, 1973), extraversion and low psychoticism (Hekmat, Khajavi, & Mehryar, 1974), and low neuroticism and psychoticism (Hekmat, Khajavi, & Mehryar, 1975).
Hogan has suggested that scores on the Empathy Scale should be related to performance as a counselor (Greif & Hogan, 1973). Some limited support for this idea was given by Sandler (1972) who found that 50 adult women who had been hired as child aides were significantly higher on Hogan's Empathy Scale than were 40 control women. Wallston and Weitz (1975) found that Hogan's Empathy Scale correlated positively with a self-report measure of acceptance of others and negatively with a self-report measure of genuineness. They found no correlation between scores on the Empathy Scale and their judge-rated empathic responses to eight, video-taped, simulated interview segments. Thus, the Hogan Empathy Scale has been validated by other self-report measures, but has not been validated against any behavioral or therapy measures of empathy. Therefore, the clinical relevance of the variable measured by the Hogan Empathy Scale is dubious. This measure may have face validity, but once again, the requirements of construct validity have not been satisfied.
CHAPTER II
RATIONALE

It is clear from the clinical definitions of empathy in the preceding chapter that the empathy process is a complex one, involving many operations on the part of the therapist. By examining empathy closely, several of the components of a therapist's behavior which together comprise the empathy process may be identified. In choosing the specific components to be examined in this study, the following criteria were used:

1. The components were taken from available clinical descriptions of the empathy process. Selection was not limited to the available operational definitions of empathy.

2. These qualities of the therapist were expected to be observable in both general, day-to-day activities and in therapeutic interactions. That is, they were to have some reasonable therapeutic relevance, but not to be limited to any esoteric behaviors or techniques of therapy.

3. The components were thought to be discrete enough to be reasonably distinguishable to a layperson. Friends and associates of the therapist should have been able to discern these qualities, as should his/her clients.

Empathy Components

The following were the hypothesized qualities of the therapist which, if operative during therapy, contribute to the process of
empathic understanding by the therapist.

1. Attention and responsiveness to others—person shows an active interest and concern for others; in conversation he/she is "with" the other person (indicated by eye contact, nodding, "um-hmm", etc.) rather than apparently looking at or thinking about something else; attends to all available cues from other person (verbal and nonverbal cues, feeling and emotion as well as the content of other's statements).

2. Ability to temporarily suspend values—respects other's ideas, beliefs; is non-judgmental; can experience and understand another person's feelings, even if contrary to his/her own value system.

3. In touch with own feelings—person is fully aware of his/her own feelings and able to separate them from other's feelings. There is a lack of projection; person does not assume other's feelings for his/her own, nor ascribe his/her own feelings to others.

4. Expression of positive regard—person conveys a sense of acceptance and personal endorsement of other people.

5. Clarity in personal communications—verbal expression is concrete, concise, articulate; language is non-esoteric, clearly understandable to others; verbal and non-verbal messages agree.

6. Strong sense of security—person is aware of who he/she is and has high degree of self-acceptance. (This allows person to freely enter another's psychological state without fear of losing his/her own identity).
Major Hypotheses

The following hypotheses were examined:

**Hypothesis 1.** Average client ratings of therapist empathy correlate positively and significantly with composite ratings of therapist empathy by professional associates.

This was an examination of the question discussed in the introduction about the relationship between therapists' behavior outside of therapy and their behavior in therapy. The professional associates' ratings of therapists were made on their empathy-related behaviors which occurred outside of the therapy setting. The client ratings were a measure of empathy-related behaviors displayed by these same therapists in therapy.

**Hypothesis 2.** Therapists' scores on the Hogan Empathy Scale correlate positively and significantly with: (a) composite ratings of therapist empathy by professional associates; (b) average client ratings of therapist empathy.

This was an assessment of the convergent validity of the Hogan Empathy Scale—a questionnaire—with two behavioral measures of empathy. The professional associates' ratings were based on non-therapy behaviors and the Hogan Scale was a self-report measure of general (non-therapy) empathic behavior. The client ratings assessed therapists' empathic behavior in therapy and Hogan has stated that therapists' scores on the Hogan Empathy Scale ought to correlate with in-therapy behavior. The Hogan Scale has not, however, been validated by measures of behavior in or out of therapy.
The Hogan Scale is ambiguous in that the factors involved in the final "empathy" score remain unspecified. The Hogan Scale was used to assess its usefulness as a measure of empathy in therapy-related settings.

**Hypothesis 3.** Therapist and client ratings of therapy outcome are positively related to therapist empathy as rated by:
(a) clients; (b) professional associates.

The first part of this hypothesis, that client-perceived empathy is positively correlated with therapeutic outcome, was basically a repetition of findings by Rogers et al. (1967). The second part of the hypothesis interjected a new proposition, that therapists' non-therapy, empathy-related behaviors are related to achieved client outcome in therapy. This hypothesis was a minor part of the overall study. There was no examination of pre- and post-outcome measures to indicate the nature or degree of change in therapy. It was a tentative look at the relationship between therapist and client perceptions of outcome and the components of empathy displayed by therapists.

**Hypothesis 4.** Using the six principal components of empathy, the ratings of therapists by clients and by professional associates possess discriminant validity.

As mentioned previously, it is important to determine whether or not the empathy measure used possesses discriminant validity. A useful method for assessing discriminant validity was devised by Campbell and Fiske (1959) in their use of a multitrait-multimethod matrix (MMM). This involves measuring
two or more different traits or qualities with two or more different tests or methods, and it was designed to control for methods variance.

In this analysis a 2 X 6 (methods X traits) matrix was used. The two methods were the client ratings and the professional associate ratings. The six traits are those listed above; each therapist was rated on the six traits by both clients and professional associates. Using client ratings as the criterion, professional associates' ratings on the components of empathy were examined for discriminant validity. An example of this matrix is demonstrated in Table 1 for three traits and two methods.

Table 1

An Example of a Multitrait-multimethod Matrix
The validity coefficients are underlined in Table 1; they represent the measurement of the same trait by two different methods. The reliability coefficients are enclosed by parentheses; these were not determined in this study where the methods were administered only once. The heterotrait-monomethod triangles are enclosed by solid lines, while the heterotrait-heteromethod triangles are enclosed by broken lines.

The validity coefficients should be significantly greater than zero as a demonstration of convergent validity. Discriminant validity for a particular trait is demonstrated when three criteria have been met. First, the validity coefficient for that trait must be greater than any coefficient lying in its row or column in the heterotrait-heteromethod triangles. This simply means that a validity value for a trait must be greater than any correlation of that variable with another variable having neither trait nor method in common. For example, the correlation of clients' and professional associates' ratings on "in touch with own feelings" should be greater than the client ratings of this trait and the associate ratings of "ability to temporarily suspend values" or "attention and responsiveness to others".

Second, the validity coefficient of that trait must be greater than any coefficient lying in its row or column in the heterotrait-monomethod triangles. This means that the correlation of the same trait using two different methods is greater than any of the correlations of one trait with other traits using the same method. For example, client and associate ratings of "in touch with own
feelings” should correlate more highly with each other than do client ratings of this trait and client ratings of either "ability to temporarily suspend values" or "attention and responsiveness to others".

Third, the same pattern of relationships must exist between traits within all heterotrait triangles of both the monomethod and heteromethod blocks. Correlations within the heterotrait triangles represent random variance and there should, therefore, be no great discrepancies in the level of correlations between traits within these triangles. For example, if client ratings of "in touch with own feelings" correlate .2 or .3 with client ratings of "ability to temporarily suspend values", then client ratings of "in touch with own feelings" ought to also correlate around .2 or .3 with client ratings of "attention and responsiveness to others".

Further explorations. An additional use of the six rating scales was made in order to assess their possible usefulness as idiographic descriptors of therapists’ personalities. This involved a separate rank-order for each subject (therapist) by each of the raters on the six empathy components. This was a fundamentally different use of the scales. The rating from one to seven was a general scaling or weighting of each component for each subject. The rank-order procedure, on the other hand, asked each rater to describe the subject on each component relative to the other five components.
CHAPTER III

METHOD

Subjects

Prior to the study, all student-therapists in the clinical psychology program at the University of Massachusetts were considered as possible subjects. An initial screening process identified as potential subjects 22 student-therapists who had taken at least one academic course from both of the faculty used as raters (see section on Raters). Each student was contacted personally, the study was explained to them and their participation was requested (see Consent Form, Appendix A). All student-therapists who were currently seeing (or had seen within the previous calendar year) at least two adult clients (16 years or older) and who agreed to participate were retained as subjects in the study. Two student-therapists had an insufficient number of clients to satisfy the criterion and eleven student-therapists declined to participate. In the final sample, a total of nine student-therapists were retained as subjects for the study.

Raters

Two sets of raters were used to rate subjects: 1) all available adult clients of each subject (therapist); 2) three professional associates of each subject (one male and one female graduate faculty member and the investigator, a male graduate student in clinical psychology).
The only selection criteria imposed by the investigator for clients was that they must have seen their therapist for a minimum of five sessions. Therapists were also asked to eliminate, based on their clinical judgment, any clients to whom participation in the study could cause some clinical harm; a total of four clients were excluded for this reason.

The two faculty members were selected on two criteria:
1) each had taught a clinical psychology core course (a required course for all clinical students) in which each of the subjects had been a class member; 2) neither faculty member had directly supervised the therapy of any of the subjects. It was believed that the common experience of the core courses would give each faculty member a roughly equivalent experiential base for rating each of the subjects. Additionally, these faculty members' ratings were based on subjects' behavior outside of therapy; if the faculty members had supervised the subjects' therapy work, it could have potentially contaminated the comparison between ratings made on in-therapy behavior (by clients) and out-of-therapy behavior (by professional associates).

The investigator's ratings were based on a variety of experiences with each of the subjects in both academic and informal settings; the experience base across each of the subjects was the least consistent of the three professional associates. It was believed, however, that such a peer rating might contribute some true variance to the composite rating of the three associates, since the faculty ratings were based primarily on academically related experiences and
the variate in question included all non-therapy behavior.

Measures

Client ratings of empathy. Client ratings of their therapist on a 1-to-7 scale (see Appendix B) of six components of empathy, yielding also a composite measure of perceived therapist empathy. Clients also described their therapist by ranking all of the qualities from one to six, from most characteristic to least characteristic.

Professional associate ratings of empathy. Personality ratings of therapists on a 1-to-7 scale (see Appendix B) of six components of empathy by three professional associates based on previous experiences with the subjects in non-therapy settings. These six components also yielded a composite measure of empathic behavior outside of therapy. Additionally, each professional associate completed a rank order of the six components for each subject.

Hogan Empathy Scale. The Hogan Empathy Scale, a self-administered, paper-and-pencil measure of empathy, completed by each subject (see Appendix C).

Therapist progress/outcome ratings. Therapist-judged assessment of treatment progress or outcome for each client based on at least five sessions with the client (see Appendix D, Therapist Form).

Client progress/outcome ratings. Client-judged assessments of treatment progress or outcome for each client based on at least
five sessions with their therapist (see Appendix D, Client Form).

Reliability of Measures

Professional associate ratings of empathy. Prior to the ratings of subjects on the six components of empathy by professional associates, a pilot study was conducted to further clarify the components and to insure interrater reliability. A manual for rating the six components on a 1-to-7 scale was devised and the three associates rated five student therapists not acting as subjects in the study. Ebel's (1951) statistic for estimating reliability of multiple raters was employed. For scales 1 through 6 in the pilot study the r was: Attention and Responsiveness to Others, .79; Ability to Temporarily Suspend Values, .85; In Touch With Own Feelings, .77; Expression of Positive Regard, .87; Clarity In Personal Communications, .67; Sense of Security, .32.

Scale 6, Sense of Security, was more carefully examined because of its low average reliability. The low reliability value resulted from a restriction in the range of scale points used by the raters (only points 2, 3 and 4 were used). Relative agreement on the ratings was good: out of 15 possible pairs of ratings, 12 were within plus or minus 1 scale point, and 3 were within plus or minus 2 scale points. After discussion between the three professional associates, some ambiguous wording on the scales was clarified and the scales were put in the final form used by the clients and associates for rating subjects. The professional associates were also encouraged to attempt to use the entire scale
range when making their ratings on study subjects.

Client ratings of empathy. It was felt that using an average of multiple client ratings of therapist behavior would add stability to the ratings of therapists' in-therapy behavior. While a minimum criteria of at least five sessions between therapist and client was imposed, the average number of sessions previous to clients' ratings was 15. Five of the therapists had two clients each as respondents and four therapists had three clients each as respondents for a total of 22 clients who responded out of 26 clients who were contacted.

Hogan Empathy Scale. The Hogan Scale of Empathy has reported test-retest reliabilities of .84 and was therefore administered only once, following the initial contact with the investigator.

Therapist and client progress/outcome ratings. As noted previously, the rating of progress/outcome was a crude measure, a 1-to-7 rating on a continuum from very unsuccessful to very successful. It was felt that the minimum of five sessions was sufficient time for therapists and clients to form a general impression of the progress of therapy to that point, and it was only this general impression that the outcome rating was intended to assess.

Procedure

Before the final pool of nine subjects was formed, the professional associates completed empathy ratings of all 22 potential subjects. All of these initial ratings as well as subsequent ratings by clients were coded with letters or numbers to keep the investigator
blind as to the identity of subjects and their corresponding scores.

Potential subjects who agreed to participate in the study were given the Hogan Empathy Scale, and arrangements were made for contacting their respective clients. Clients were informed by their therapists that they would be asked to fill out a questionnaire for a research project; clients received and returned the completed ratings by mail. Clients were informed that their ratings were for research purposes only, were not to be used in any evaluation of them or their therapist's performance, and that their ratings would not be seen by their therapist.

Therapists were informed that the study was a cross-validation study of several measures of empathy. They were not informed of the specific qualities they were being rated on until after clients had completed their ratings. Therapists completed their ratings of the progress/outcome of therapy when the clients received their questionnaires in the mail.
CHAPTER IV
RESULTS

Data analysis. Statistical analysis of the data from this study was primarily correlational in nature; the Pearson product moment correlation coefficient was employed. All tests of significance were evaluated at alpha = .05. Ebel's (1951) statistic was used to compute the reliability of average ratings for multiple raters for interval scaled data. Guilford's (1954) formula was used to compute the reliability of average ratings of multiple raters for rank-ordered data.

In analyzing the concordance among the professional associates, interrater agreement was also computed. Tinsley and Weiss (1975) distinguish between interrater reliability and interrater agreement. Interrater agreement exists when raters tend to assign the same absolute values to the persons being rated. On the other hand, high reliability indicates that ratings closely resemble each other in the degree to which they deviate from their respective means. Thus, interrater reliability is a measure of the relative arrangement of one set of ratings to another set of ratings. As long as the relationship between persons being rated remains the same across different judges, the absolute values or scores assigned by raters do not have to be the same to achieve high reliability.

The optimal situation, of course, is for ratings to have both high agreement and high reliability. It has been shown that ratings may have high reliability and low agreement. Another possibility exists where ratings have high agreement and low reliability. This
latter situation may occur when the range of ratings used by raters is restricted due to similarity in persons being rated or improper use of the scale by raters. To completely assess the homogeneity of sets of ratings it is necessary, therefore, to examine both interrater reliability and interrater agreement.

Reliability and Agreement Between Professional Associates

For the analysis of interrater reliability and agreement between the three professional associates only, the original subject pool of 22 potential subjects was used. This treatment of the ratings was done to assess the reliability and validity of the empathy rating instrument as used by the associates. It was believed that this part of the data analysis was more meaningful than if only the final nine subjects had been used because final completion of the study (i.e., agreeing to participate) constituted an artificial sampling of available subjects. Otherwise, willingness or ability to cooperate in the study would have served as an intervening variable, and this was not a desired treatment condition.

Ratings made by professional associates were analyzed by using a mult trait-multimethod matrix with three methods (raters) and six traits (empathy components) (see Table 2).
Table 2

Multitrait-multimethod matrix:
Professional Associates × Professional Associates

<table>
<thead>
<tr>
<th>TRAITS</th>
<th>A1</th>
<th>A2</th>
<th>A3</th>
<th>A4</th>
<th>A5</th>
<th>A6</th>
<th>B1</th>
<th>B2</th>
<th>B3</th>
<th>B4</th>
<th>B5</th>
<th>B6</th>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
<th>C5</th>
<th>C6</th>
</tr>
</thead>
<tbody>
<tr>
<td>METHOD A (rater)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1</td>
<td>()</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A2</td>
<td>44*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A3</td>
<td>00*20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A4</td>
<td>51*24 15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A5</td>
<td>08 36<em>40</em>05</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A6</td>
<td>31 20 19 27 54*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>METHOD B (rater)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>52*22 02 40 25 23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B2</td>
<td>44<em>63</em>33 40 13 18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B3</td>
<td>10<em>19 43</em>01*14 10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B4</td>
<td>40 26 00 70*13 29</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B5</td>
<td>07 23<em>13</em>19 36 07</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B6</td>
<td>13 17 12<em>07 61</em>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>METHOD C (rater)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C1</td>
<td>60<em>16 14 35</em>06 42*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C2</td>
<td>61<em>42</em>05 31 10 27</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C3</td>
<td>21<em>12 38 04</em>02 24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C4</td>
<td>69<em>18 09 35</em>02 39</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C5</td>
<td>21 26<em>12 05 39 48</em></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>C6</td>
<td>28 02 04 13 00 61*</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* p < .05
Values given are in hundredths
Values in the validity diagonals are underlined
The monomethod blocks (Rater A X Rater A, etc.) reveal the degree to which each of the three associates discriminated among the six traits when making their ratings. Raters A and B discriminated fairly well; for each rater, in only 3 cases out of 15, their ratings on one trait were significantly correlated with their ratings on another trait. Rater C, however, did not discriminate well among the six traits; in 14 cases out of 15, Rater C's ratings on one trait were significantly correlated with Rater C's ratings on another trait. For Rater C, some kind of halo effect was operating; this rater tended to rate a subject either high on all six traits or low on all six traits. Rater C was one of the two faculty members, so the difference could not be accounted for by whether the rater was a student or faculty member. There was no data to suggest why the halo effect was operating for Rater C and not for the other two raters.

Table 3 summarizes the information from Table 2 necessary for assessing the convergent and discriminant validity of the six scales for the three professional associates.
### Table 3

Convergent and Discriminant Validity Between Professional Associates *

<table>
<thead>
<tr>
<th>Scale</th>
<th>Rater A × B</th>
<th>Rater A × C</th>
<th>Rater B × C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A</td>
<td>B</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>.56 (0) [0] [0]</td>
<td>.60 (2) [0] [4]</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>.63 (0) [0] [0]</td>
<td>.44 (1) [0] [3]</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>.43 (0) [0] [1]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>.70 (0) [0] [0]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>.73 (0) [0] [0]</td>
<td>.61 (0) [0] [3]</td>
<td></td>
</tr>
</tbody>
</table>

* All correlations shown are significant, p < .05

Enclosed numbers adjacent to each validity coefficient indicate the number of coefficients exceeding that validity coefficient in its respective heteromethod block ( ) and monomethod blocks [ ].

The ratings between Rater A and Rater B demonstrated fairly good convergent and discriminant validity for scales 1, 2, 3, 4 and 6. Each of these validity coefficients was significant—a demonstration of convergent validity. Discriminant validity was also demonstrated since the validity coefficients exceeded the other correlations in their respective row or column in both the monomethod and heteromethod blocks. There was one exception for scale 3, where the validity coefficient of .43 was exceeded by the .49 correlation of scales 1 and 3 for Rater B.
Between Raters A and C, scales 1, 2, and 6 demonstrated convergent validity; their discriminant validity was not clearly demonstrated, however. The validity values were generally higher than the heterotrait values in the heteromethod block; they were not, however, consistently higher than the heterotrait values in the monomethod block of Rater C.

Between Raters B and C, only scale 3 demonstrated convergent validity. It also clearly demonstrated discriminant validity against the other five scales.

In general, scales 1, 2, 3, and 6 demonstrated relatively good convergent and discriminant validity for the three professional associates. Almost all of the methods variance in scales 1, 2 and 6 was accounted for by the lack of discrimination between traits made by Rater C. Scale 4 demonstrated some evidence of convergent and discriminant validity; the scale had very good evidence of validity between Raters A and B, but not between either Raters A or B and Rater C. The validity coefficient for scale 5 did not reach significance between any of the raters.

The average reliabilities and agreements between the professional associates for the six scales is presented in Table 4.
Table 4
Average Reliabilities and Agreements Between Professional Associates

<table>
<thead>
<tr>
<th>Scale</th>
<th>Ebel's r (Reliability of avg. of 3 raters)</th>
<th>Agreement (Percentage of agreements within ± 1 pt.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.93</td>
<td>80 %</td>
</tr>
<tr>
<td>2</td>
<td>.87</td>
<td>71 %</td>
</tr>
<tr>
<td>3</td>
<td>.73</td>
<td>65 %</td>
</tr>
<tr>
<td>4</td>
<td>.71</td>
<td>79 %</td>
</tr>
<tr>
<td>5</td>
<td>.56</td>
<td>52 %</td>
</tr>
<tr>
<td>6</td>
<td>.78</td>
<td>65 %</td>
</tr>
</tbody>
</table>

Ebel's statistic was used to compute average reliabilities. Correct agreement was defined as those ratings which were within plus or minus one point of each other on the seven point scale; the agreement score for each scale was the percentage of correct agreements out of all ratings made. The reliability and agreement figures were consistent with the validity conclusions made from the MMM analysis. Scales 1, 2, 3, 4 and 6 had relatively good average reliability and agreement for the professional associates; scale 5 failed to demonstrate good reliability or agreement.

Major Hypotheses

For the remaining statistical analyses involving the four main hypotheses, only the data for the final study subjects was used, with \( N = 9 \). The composite empathy score for subjects was
computed by summing the values assigned by raters across all six scales. Averages of clients' composite ratings were computed to account for the unequal number of clients for each subject.

**Hypothesis 1.** Hypothesis 1, that clients' composite empathy ratings correlate positively and significantly with professional associates' empathy ratings, was not supported (see Table 5).

**Table 5**

<table>
<thead>
<tr>
<th>Average Composite Empathy Ratings by Clients</th>
<th>Composite Empathy Ratings by Professional Associates (Combined)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.50</td>
</tr>
<tr>
<td>Composite Empathy by:</td>
<td></td>
</tr>
<tr>
<td>Rater (Associate) A</td>
<td>.36</td>
</tr>
<tr>
<td>Rater (Associate) B</td>
<td>.73*</td>
</tr>
<tr>
<td>Rater (Associate) C</td>
<td>.25</td>
</tr>
</tbody>
</table>

*<p < .05

Although the ratings made by Rater B alone were significantly correlated with average client ratings, the ratings by all associates combined were not significant. Rater B was one of two faculty members, so the difference could not be accounted for by whether the rater was a student or faculty member. There was no data to suggest why the ratings by Rater B were significantly correlated with client
ratings, while the ratings by Rater A and Rater C were not.

**Hypothesis 2.** Hypothesis 2, that subjects' scores on the Hogan Empathy Scale correlate positively and significantly with empathy ratings by clients and by professional associates was not supported. Subjects' scores on the Hogan Empathy Scale had Pearson r correlations of .43 (ns) with professional associates' composite empathy ratings and .59 (ns) with average client empathy ratings.

**Hypothesis 3.** Hypothesis 3, that therapist and client ratings of therapy outcome correlate positively and significantly with empathy ratings by clients and by professional associates was not supported (see Table 6).

<table>
<thead>
<tr>
<th>Progress/Outcome Ratings by Clients</th>
<th>Average Composite Empathy Ratings by Clients</th>
<th>Composite Empathy Ratings by Professional Associates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.27</td>
<td>-.32</td>
</tr>
<tr>
<td></td>
<td>.04</td>
<td>.40</td>
</tr>
</tbody>
</table>

None of the correlations between rated empathy and rated progress/outcome were significant. The Pearson r correlation between therapists' (subjects') ratings of progress/outcome and clients' ratings of progress/outcome was .10. This low correlation
was due primarily to the restricted range of progress/outcome scores assigned by therapists and clients; ratings were heavily skewed toward the upper (positive) end of the scale. In terms of agreement, 62% of the therapist and client ratings were within plus or minus one point on the seven point progress/outcome scale; 100% of the ratings were within plus or minus two scale points.

Hypothesis 4. This was an examination of the validity of the professional associates' ratings of subjects against the ratings made by clients. This matrix is presented in Table 7 with six traits (empathy components) and two methods (composite professional associates' ratings and average clients' ratings).
### Table 7
Multitrait-multimethod Matrix: Professional Associates X Clients

<table>
<thead>
<tr>
<th>TRAITS</th>
<th>METHOD C (Clients)</th>
<th>METHOD A (Associates)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C1  C2  C3  C4  C5  C6</td>
<td>A1  A2  A3  A4  A5  A6</td>
</tr>
<tr>
<td>C1</td>
<td>()</td>
<td>()</td>
</tr>
<tr>
<td>C2</td>
<td>59</td>
<td>()</td>
</tr>
<tr>
<td>C3</td>
<td>51  35 ()</td>
<td>()</td>
</tr>
<tr>
<td>C4</td>
<td>63  70  63  63  63  ()</td>
<td>69  69  69  69  69  ()</td>
</tr>
<tr>
<td>C5</td>
<td>63  76  18  73  73  ()</td>
<td>63  73  73  73  73  ()</td>
</tr>
<tr>
<td>C6</td>
<td>73  72  54  93  81  ()</td>
<td>73  73  73  73  73  ()</td>
</tr>
<tr>
<td>A1</td>
<td>03  11  21  08  30  22</td>
<td>()</td>
</tr>
<tr>
<td>A2</td>
<td>59  61  25  38  45  45</td>
<td>45 ()</td>
</tr>
<tr>
<td>A3</td>
<td>34  35  64  49  57  59</td>
<td>65  21 ()</td>
</tr>
<tr>
<td>A4</td>
<td>23  23  32  29  51  37</td>
<td>93  51  77 ()</td>
</tr>
<tr>
<td>A5</td>
<td>21  14  08  17  28  46</td>
<td>39  43  00  27 ()</td>
</tr>
<tr>
<td>A6</td>
<td>30  30  11  18  69  24</td>
<td>72  52  46  82  42 ()</td>
</tr>
</tbody>
</table>

* p < .05
Values given are in hundredths
Values in the validity diagonals are underlined

None of the validity coefficients for the six scales were significant. By examining the monomethod block for clients, it will be noticed that clients did not discriminate well among the six traits. Nine cases out of 15 heterotrait correlations were significant. The professional associates as a group were better able to discriminate; for the associates, only four out of 15 heterotrait correlations were significant.
A strong contributing factor to the low reliabilities was the lack of discrimination made by clients; their ratings were strongly skewed to the upper end of the scale, causing a great restriction in the range of scores assigned to subjects. A total of 54 average-scores were given to subjects by clients (9 subjects × 6 traits); in only 7 of these 54 scores was there an average score below 5 on the 7-point scale. Because of their restricted range, clients' ratings contributed very little variance to the final reliability estimates.

Agreement scores between clients and associates were not computed because the severe restriction in the range of client scores made such an analysis meaningless.

Nonparametric analysis of the empathy rating scales. The final analysis involved the rank-order of the six components for each subject made by professional associates and by clients (see Table 8).
Table 8

Rank-order Correlations of Empathy Components by Professional Associates and Clients

<table>
<thead>
<tr>
<th>Subject</th>
<th>Average Rank-order Correlation of 3 Professional Associates</th>
<th>Clients' Average Rank-order</th>
<th>Professional Associates' Average Rank-order</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.01</td>
<td></td>
<td>-.33</td>
</tr>
<tr>
<td>2</td>
<td>.18</td>
<td></td>
<td>.30</td>
</tr>
<tr>
<td>3</td>
<td>.50</td>
<td></td>
<td>.02</td>
</tr>
<tr>
<td>4</td>
<td>.62</td>
<td></td>
<td>.48</td>
</tr>
<tr>
<td>5</td>
<td>.10</td>
<td></td>
<td>.11</td>
</tr>
<tr>
<td>6</td>
<td>.09</td>
<td></td>
<td>.91*</td>
</tr>
<tr>
<td>7</td>
<td>.73</td>
<td></td>
<td>.65</td>
</tr>
<tr>
<td>8</td>
<td>.85</td>
<td></td>
<td>.31</td>
</tr>
<tr>
<td>9</td>
<td>.30</td>
<td></td>
<td>.71*</td>
</tr>
</tbody>
</table>

*p < .05

The numbers in the left-hand column represent the average rank-order correlations of the three professional associates for each subject (Guilford, 1954, p. 397). The professional associates were not able to make consistent agreements in their rank-order descriptions of each subject.

The numbers in the right-hand column represent the correlations between the average rank-order descriptions of subjects made by professional associates and by clients. To compute this correlation, each component for each therapist was assigned the value of its
average ranking by clients. A second value was assigned to each component from the average of professional associates' rankings. Because of the averaging procedure, some components had tied values assigned to them. Pearson's $r$ was computed for the two sets of obtained values. The obtained $r$ values between average client rankings and average professional associates rankings were generally low. There was significant agreement for subjects six and nine; the correlations for the other seven subjects were not significant.
CHAPTER V
DISCUSSION

The results of this study add little to the theoretical issues relating to empathy. None of the four major hypotheses was confirmed, and because of inadequacies in the study itself, the results remain ambiguous. Results will be discussed relative to problems in the design and execution of the study and only modest speculation can be offered about any substantive theoretical issues. The feasibility of using the empathy component scales in future research will also be discussed.

Problems In Design: Major Hypotheses

The most obvious shortcoming in this study was the small sample size. A canon of inferential statistics demands that for strong inferences to be made about a given population (in this case, psychotherapists), a sufficiently large sample from that population must be examined on the variates in question. The larger the sample size, the less likely it is that one or more chance variables will affect the outcome; accordingly, confidence levels for rejecting a null hypothesis become exceedingly more stringent as the sample size decreases. The requirement of a suitable sample size was obviously not met in this study where the final sample had only 9 subjects. Although most of the correlations between the variates in question were in the predicted direction, the results cannot be said to be attributable to anything more than chance.
**Hypothesis 1.** The results did not confirm the notion put forth by Truax and Carkhuff (1967), Bergin and Strupp (1972) and Rogers (1975) that the ability to be empathic signifies a general personality characteristic of a therapist. There was minimal support for the idea that a therapist's behavior outside of therapy is related to how s/he behaves in therapy; one of the professional associate's ratings correlated significantly with clients' ratings. Ratings by the three professional associates combined, when compared to clients' ratings, were correlated in the predicted direction although the correlations were not significant. This hypothesis would warrant further investigation.

**Hypothesis 2.** There was little support for the supposition by Hogan (Greif & Hogan, 1973) that scores on the Hogan Empathy Scale predict a person's ability to function as an empathic therapist. The information obtained regarding the Hogan Empathy Scale was insufficient to make a judgment about its usefulness for rating therapist empathy. The small sample size was again a primary factor in the lack of clear results regarding this scale. Additional testing needs to be done to confirm the construct validity of the Hogan Scale.

**Hypothesis 3.** There was no support for Rogers' (1957) belief that client-perceived empathy is an essential element of success in therapy. One interesting correlation here was the negative correlation between progress/outcome ratings by clients and professional associates' ratings of therapists' empathy behavior outside of therapy. Although the correlation was not significant, it does invite the
further exploration of this previously unexplored proposition—that a therapist's empathy-related behavior outside of therapy is related to the outcome their clients achieve in therapy.

It also raises a question of whether or not a client's rating of progress/outcome is an adequate measure of what happens in therapy. In general, the therapist and client ratings of outcome were not adequately assessed for their usefulness. The self-rating by therapists and clients still may be an important source of information for evaluating outcome. In future use, a pre- and post-test comparison of ratings would be useful. It might also be useful to institute some objective ratings of outcome (standardized test scores, trained observers' ratings); this, however, is actually an entire field of research in itself. Just as there is little agreement among researchers regarding the best way to measure empathy, so also is there dissension about the best way to assess outcome. It is clear that the measurement techniques for both empathy and therapeutic outcome need to be better refined; hopefully then, something more specific can be said about how therapist qualities interact with client qualities to achieve certain kinds of outcome.

Hypothesis 4. The construct validity for the six empathy components was not clearly demonstrated. Using Rogers' stipulation that the client's perceptions are crucial, the client ratings served as the validity criterion; the ratings made by professional associates demonstrated neither convergent nor discriminant validity against client ratings. There was some evidence of construct validity for the scales between professional associates alone.
The lack of clear construct validity for the six scales further adds to the indeterminate nature of the results for Hypotheses 1, 2 and 3. That is, it cannot be said that the empathy component scales were clearly measuring what they purported to measure. As mentioned, a major factor for this failure was the small sample size. In future use, it would be essential to use a large sample size to assess the scales for construct validity before making substantive conclusions based on the scales.

Rank-order description of therapists. The rank-order procedure used for the six empathy components does not appear to be an effective way to use the scales. The idea behind this procedure was to force a discrimination by the raters between the six traits. This idea of ordering traits to describe a person was taken from the Q-sort method devised by Block (1961). With only six traits, however, the difference between each rank is much more critical than in a typical Q-sort where 100 items are sorted into 10 separate piles or rankings. It does not appear that the procedure of rank-ordering is a useful way to achieve discriminant validity for the six empathy components.

Additional Problems in Design

Design problems are compounded by the lack of experimental controls that researchers may typically impose when examining psychotherapy. There are, of course, important clinical safeguards which are necessary to maintain the personal well-being of both therapist and client. However, difficulties can exist
because the researcher often has little or no input prior to the establishment of norms or conditions for therapy which might conceivably facilitate the conduct of research. Research questions about therapy must often be asked in a limited fashion; the researcher may step into a clinic-therapist-client system in which both explicit and implicit rules or expectations are fairly well established. Such a circumstance automatically limits the scope of the questions which may be asked and the results which may be expected.

Such was the case in this study where the investigator was able to take little initiative in establishing the conditions of the research; the ideal design was compromised by already existing norms of the social milieu in which the study was conducted. The atmosphere in the clinic where the research was conducted was not conducive to the research; there was no expectation that therapists and clients would take part in any research projects. The investigator was sometimes in the position of trying to persuade student-therapists to participate as subjects. While there is no way of knowing the experimental effects of the investigator's ability or inability to persuade subjects to participate, it probably did have some effect on the make-up of the final subject pool. Difficulty in recruiting subjects was an inhibiting factor in the design of this study.

Problems with Measures and Procedures. Although it is difficult, because of the small sample, to identify clear and consistent deficiencies in the procedures and instruments employed, some
general observations can be offered. One problem already noted in the Results section was the restriction in the range of ratings made by clients, with most ratings at the high (positive) end of the scales. One possible explanation for this might be that in the recruitment of subjects, only those therapists who were personally secure or successful in therapy were willing to participate in the study. In that case, the self-selection process which operated for subjects yielded a sample from only the high end of the distribution for the variates in question. A restricted range phenomenon would result from such a sampling error. Although this seems to be one plausible explanation, it is not likely that this factor alone accounts for the restricted scores and some other possible reasons may also be suggested.

The occurrence of "yeasaying" is a recognized response bias for clients wherein they tend to respond in an agreeable, positive manner to any kind of evaluative questionnaire (Couch & Keniston, 1960). It is possible that this response bias was operative in this study, even though clients were told that their ratings were strictly for non-evaluative research purposes. It also seems that a client would want to, as a matter of potential personal benefit, perceive their therapist as a healthy, helpful individual. Most clients are hopeful of receiving some benefits from therapy and this hopefulness alone could predispose them toward seeing any therapist as an empathic individual.

The response set of the clients is also a factor when it is contrasted with the response set of the professional associates who
rated the student-therapists. The seven point rating scales for empathy components imply that, to some extent, the assigned number is arrived at by comparing the subject with other persons in general known by the rater. It is possible that the general comparison group known by clients was different than the comparison group known by the professional associates. Clients rated subjects very high, and they may well have been very high on these components relative to other acquaintances of the clients.

Another possible explanation is that clients and professional associates were not operating with exactly the same amount of experience relative to the scales themselves. The associates were fairly sophisticated in the use of psychological terminology and had the benefit of practice with the scales in the pilot study; clients were exposed to the scales only once.

Conclusions and Suggestions for Future Research

The key problem operationally was that clients and one of the professional associates were not discriminating well among the six scales. In future use, some modification of the scales might facilitate better discrimination by raters. Given that clients may tend to rate their therapists in a positive fashion, it might be useful to modify the "negative" or low end of each scale. If the low ends of each scale were less pejorative, raters might be more inclined to use scale points 1 through 4 more often. This would have the effect of increasing discrimination by widening the range of scores assigned by raters. The danger in such a revision,
of course, would be to eliminate those instances (presumably rare) in which a client would actually rate his/her therapist in a negative fashion.

It might also be beneficial if clients could be given some demonstration of and practice in the use of the scales by some impartial trainer. Through some short training session, there is a likelihood that clients would have a better understanding of the scales and would be better able to make discriminations among the scales.

Even following such a training session for clients, there would remain the question of whether or not clients' subjective evaluations of their therapists are accurate. An additional check on validity could be instituted by having trained observers rate therapists' behavior on the empathy components, based on observations of actual therapy sessions. Trained observers' ratings of therapists' behavior could then be analyzed with the multitrait-multimethod matrix against clients' ratings and/or associates' ratings; this would be an important check on the construct validity of the scales. The use of trained raters is similar, of course, to the method used for the Truax Accurate Empathy Scale, although the Truax scale has failed to demonstrate construct validity in spite of its popularity in therapy research.

The difficulty in achieving adequate raters' discrimination among traits is, of course, largely a function of the traits themselves. All of the traits were hypothesized components of the same general characteristic of empathic ability; therefore, it
would be reasonable to expect a highly empathic person to have somewhat high levels of all of the contributing components. It is obviously a difficult task to take a complex theoretical construct with subtle differences in its component parts and to then operationalize them in relatively concise and easily understood rating scales. It may indeed be a near-impossible task without extensive training for the raters who will use the scales. However, it is certainly a worthwhile endeavor to attempt greater discrimination among traits through refinement of the instruments and procedures used to measure those traits. For the construct of empathy at least, the alternative has been to rely on unproven measures which give only homogenized glimpses of the construct. To continue to rely on these unsatisfactory measures will ultimately frustrate any attempt to truly understand the empathic process.
References


Conrey, A. L., Backer, T. E., & Glaser, E. M. (Eds.). A sourcebook for


Truax, C. B. Therapist empathy, warmth, and genuineness and patient personality change in group psychotherapy: a comparison between interaction unit measures, time sample measures, and patient perception measures. Journal of Clinical Psychology, 1966, 22, 225-229. (b)


Appendix A

Subject Consent Form

This study will involve the participation of therapists and their respective clients, currently working in the PSC. It is essentially a cross-validation study of several different measures of therapists' behavior, both in and outside of therapy. The study will attempt to assess whether or not certain behaviors related to empathy generalize across a number of situations. This necessitates gathering information about your behavior from a number of sources, including: 1) your responses to a self-report questionnaire; 2) clients' ratings based on experience with you in therapy; 3) faculty members ratings based on experiences with you in academic settings (instructors from past clinical core courses--Bonnie Strickland and Norm Watt); 4) peer ratings (I will be making these ratings). These ratings will be inter-correlated to determine similarities or differences in behavior across different settings. Ratings will be used strictly for research purposes and will not be part of any evaluative process; all ratings will be held strictly confidential.

As a Subject, you will be asked to do the following: 1) fill out a standardized 64-item, true-false questionnaire; 2) rate on a 1 to 7 scale the progress or outcome of therapy with each of your clients.

Your clients will be asked to do the following: 1) make ratings of your behavior in several areas; 2) rate progress or outcome in therapy on the same 1 to 7 scale used by you, their therapist. Your clients will be informed that their ratings are confidential (that you, their therapist, will not see the ratings), are for research purposes only and will not be used in any evaluation of your work as a therapist. You in turn are requested not to question your clients regarding the nature of the ratings or their responses.

If clients initiate discussion of their ratings with you, and you feel that this discussion is clinically relevant or important to your therapeutic relationship with that client, you are requested to keep such discussion to the minimum level you deem necessary. It seems likely, for example, that some discussion of progress may be a natural part of the therapy process. The essential point here is for clients not to feel in any way pressured into revealing their responses to the ratings scales. If a client does voluntarily share with you any of the content of the ratings scales, you are requested not to share this information with others.

At the conclusion of the study you may see the final aggregate results as presented in the Results and Discussion sections of the study, including the specific components of behavior on which you were rated by clients, faculty and peer(s). Individual ratings or scores will not be available to you or others to view.

You may withdraw consent and discontinue your participation in the study at any time.

I understand and agree to the above conditions and agree to participate in the study.

(Signature of Subject)
Appendix B

Rating Scales for Empathy Components

On the following pages you will be asked to answer some questions about your therapist and your experiences in therapy. Please answer all sections completely and be as honest as you can. Please rate your own, individual perceptions, based on your experience with the therapist.

Following are six personal qualities or characteristics. Each of these characteristics is to be rated on a separate scale from one to seven (low to high). Each quality you are to rate is listed in capital letters and is followed by a brief definition of that quality.

To assist you in making the ratings, further description has been added to points one (low), four (medium), seven (high). These are hypothetical descriptions of a person lying at that value on the scale. Not all of the examples at these points will necessarily apply to every person at that value. These are only listed as suggestions to help you make the ratings.

Although points two, three, five, and six are not specifically described, they may also be used in the ratings. The primary quality you are to rate is described at the top of the scale; it is to be rated anywhere on a continuous scale from one to seven (low to high).

Please rate your therapist on each quality by circling the appropriate number on the scale at the bottom of each page.
ATTENTION AND RESPONSIVENESS TO OTHERS -- how emotionally sensitive is this person in interactions with other people.

**Level 1.** Person has little awareness of others' needs or concerns; seems to dislike extended personal encounters; is unable to follow your verbal expressions; doesn't follow the content of what you say (e.g., may often or unexpectedly change the subject); does not notice or attend to emotions expressed by others.

**Level 2.**

**Level 3.**

**Level 4.** Person seems aware of others' needs or concerns; responds primarily to obvious or specific requests; shows moderate spontaneous helpfulness or responsiveness; follows the content of what you say moderately well; shows moderate response to emotions expressed by others; may attend more readily to emotions which are positive/pleasant/happy than to emotions which are negative/painful/angry.

**Level 5.**

**Level 6.**

**Level 7.** Person is highly sensitive to others' needs or concerns; easily follows the content of what you say and is "with you" in conversations (shown by eye contact, nodding etc.); attends to both verbal and nonverbal messages; offers attention and concern spontaneously; is very mindful of others' emotions; recognizes and responds to negative/painful/angry emotions as well as to positive/pleasant/happy ones.
ABILITY TO TEMPORARILY SUSPEND VALUES -- how open-minded and non-judgmental is this person; is s/he insulated, shortsighted, intolerant or impartial, broadminded, tolerant?

Level 1. Person's belief system is rigid; can view situations from only one perspective; views others in an evaluative (good/bad) manner; scorns or criticizes ideas contrary to her/his own.

Level 2.

Level 3.

Level 4. Person is at least minimally aware of alternate or opposing points of view; does little open criticizing of contrary ideas; view of others is not limited to evaluative (good/bad) dimension; may make some effort to assume others' points of view but has trouble doing so; holds fairly strongly to values/beliefs if questioned or confronted by others; makes occasional judgmental comments about others.

Level 5.

Level 6.

Level 7. Person is fully aware of alternative points of view; can easily assume another person's position on an issue; occasionally modifies own beliefs or value system; avoids any direct or implied judgmental statements about others; doesn't hold unreasonably firmly to own views in the face of confrontation.
IN TOUCH WITH OWN FEELINGS -- how openly does this person talk about the feelings or emotions s/he is experiencing?

Level 1. Person expresses little emotion; emotion is very flat or restricted; person avoids making comments like "I feel . . ." or "that's how I feel"; nonverbal expressions are blank, emotionless.

Level 2.

Level 3.

Level 4. The range (type) of different emotions s/he expresses is moderate; a moderate number of self-referent statements are made about very obvious feelings being experienced; self-referent statements deal primarily with positive or negative feelings but not both; some nonverbal cues of emotions (smiling, frowning, laughter, blushing etc.) are evident but are not readily expressed verbally.

Level 5.

Level 6.

Level 7. Person is fully aware of moment-to-moment feelings and expresses them openly ("I feel . . ."); nonverbal emotional cues are often or usually accompanied by verbal expressions of these emotions; person expresses both positive and negative emotions.
EXPRESSION OF POSITIVE REGARD -- how effectively does this person convey a sense of acceptance and personal endorsement of other people?

Level 1. Person typically relates to others in hostile, rejecting ways, as if they were not liked; may come across as contemptuous or competitive; others get the distinct impression of not being appreciated.

Level 2.

Level 3.

Level 4. Person typically relates to others with indifference or alternately expresses rejection and endorsement; may come across as consistently aloof, as if all others can be taken or left, or may polarize people—turning some on and others off.

Level 5.

Level 6.

Level 7. Person typically relates to others with warmth, acceptance and a clear sense of positive regard; consistently leaves people with a feeling of being prized and endorsed personally.
CLARITY IN PERSONAL COMMUNICATIONS -- how clear are the statements made by this person; do verbal expressions seem consistent with apparent feelings?

Level 1. Person may use a lot of unnecessary words or "ramble on" at times; uses jargon or non-specific language to express self; you consistently need to ask for clarification of her/his statements; verbal and nonverbal messages do not agree.

Level 2.

Level 3.

Level 4. Person has good facility for talking about ideas or things which are real, substantial, concrete; has occasional difficulty talking about things which are abstract, intangible, subtle; you may have to occasionally ask for clarification of statements, but following this questioning the person can usually say things more clearly.

Level 5.

Level 6.

Level 7. Person is verbally concise, organized, articulate and non-esoteric; the meaning of messages is almost always clear; you seldom need to ask for clarification of what is said; verbal and nonverbal messages agree.
SENSE OF SECURITY -- what is the emotional well-being of this person; how does s/he feel about her or him self; to what degree is s/he either secure, content, solid, self-possessed or defensive, insecure, unconfident, discontented.

Level 1. Person is insecure, defensive; expresses a lot of self-doubts and may need a great deal of reassurance from others; seems discontent with self and may try to emulate or copy others.

Level 2.

Level 3.

Level 4. Moderate degree of self-acceptance is present; has some self-awareness, but this self-knowledge may be tentative or "shaky"; may become defensive or feel personally attacked by direct criticisms, suggestions or confrontations from others; may need occasional reassurance.

Level 5.

Level 6.

Level 7. Person is fully aware of who s/he is; has a high degree of self-acceptance and a strong sense of emotional well-being; can accept others' criticisms or suggestions without feeling personally attacked; seldom needs personal reassurances from others.
Using the same six personal qualities just described, describe your therapist by ranking all of the qualities from one to six, from most characteristic to least characteristic. For example, if EXPRESSION OF POSITIVE REGARD is the most characteristic of your therapist, that would be labeled number 1. If the next most characteristic quality is CLARITY IN PERSONAL COMMUNICATIONS, that would be number 2, and so on. Rate all six qualities in this way—no tied rankings are allowed.

Rank

___ ATTENTION AND RESPONSIVENESS TO OTHERS
___ ABILITY TO TEMPORARILY SUSPEND VALUES
___ IN TOUCH WITH OWN FEELINGS
___ EXPRESSION OF POSITIVE REGARD
___ CLARITY IN PERSONAL COMMUNICATIONS
___ SENSE OF SECURITY
Appendix C

Hogan Scale of Empathy

1. A person needs to "show off" a little now and then.
2. I liked "Alice in Wonderland" by Lewis Carroll.
3. Clever, sarcastic people make me feel very uncomfortable.
4. I usually take an active part in the entertainment at parties.
5. I feel sure that there is only one true religion.
6. I am afraid of deep water.
7. I must admit I often try to get my own way regardless of what others may want.
8. I have at one time or another in my life tried my hand at writing poetry.
9. Most of the arguments or quarrels I get into are over matters of principle.
10. I would like the job of a foreign correspondent for a newspaper.
11. People today have forgotten how to feel properly ashamed of themselves.
12. I prefer a shower to a bathtub.
13. I always try to consider the other fellow's feelings before I do something.
14. I usually don't like to talk much unless I am with people I know very well.
15. I can remember "playing sick" to get out of something.
16. I like to keep people guessing what I'm going to do next.
17. Before I do something I try to consider how my friends will react to it.
18. I like to talk before groups of people.
19. When a man is with a woman he is usually thinking about things related to her sex.

20. Only a fool would try to change our American way of life.

21. My parents were always very strict and stern with me.

22. Sometimes I rather enjoy going against the rules and doing things I'm not supposed to.

23. I think I would like to belong to a singing club.

24. I think I am usually a leader in my group.

25. I like to have a place for everything and everything in its place.

26. I don't like to work on a problem unless there is the possibility of coming out with a clear-cut and unambiguous answer.

27. It bothers me when something unexpected interrupts my daily routine.

28. I have a natural talent for influencing people.

29. I don't really care whether people like me or dislike me.

30. The trouble with many people is that they don't take things seriously enough.

31. It is hard for me just to sit still and relax.

32. Once in a while I think of things too bad to talk about.

33. I feel that it is certainly best to keep my mouth shut when I'm in trouble.

34. I am a good mixer.

35. I am an important person.

36. I like poetry.

37. My feelings are not easily hurt.

38. I have met problems so full of possibilities that I have been unable to make up my mind about them.
39. Often I can't understand why I have been so cross and grouchy.
40. What others think of me does not bother me.
41. I would like to be a journalist.
42. I like to talk about sex.
43. My way of doing things is apt to be misunderstood by others.
44. Sometimes without any reason or even when things are going wrong I feel excitedly happy, "on top of the world."
45. I like to be with a crowd who play jokes on one another.
46. My mother or father often made me obey even when I thought it was unreasonable.
47. I easily become impatient with people.
48. Sometimes I enjoy hurting persons I love.
49. I have sometimes felt that difficulties were piling up so high that I could not overcome them.
50. I am apt to pass up something I want to do when others feel that it isn't worth doing.
51. People have often misunderstood my intentions when I was trying to put them right and be helpful.
52. I am usually calm and not easily upset.
53. I would certainly enjoy beating a crook at his own game.
54. I am often so annoyed when someone tries to get ahead of me in a line of people that I speak to him about it.
55. I used to like hopscotch.
56. I have never been made especially nervous over trouble that any members of my family have gotten into.
57. I frequently undertake more than I can accomplish.
58. I enjoy the company of strong-willed people.
59. Disobedience to the government is never justified.
60. It is the duty of a citizen to support his country, right or wrong.
61. I have seen some things so sad that I almost felt like crying.
62. I have a pretty clear idea of what I would try to impart to my students if I were a teacher.
63. As a rule I have little difficulty in "putting myself into other peoples' shoes."
64. I am usually rather short-tempered with people who come around and bother me with foolish questions.
Appendix D

Progress/Outcome Scale: Therapist Form

On a scale from one to seven rate the outcome of therapy (or progress to this time). Please circle the appropriate number.

How successful have you been in helping the client to achieve his/her goals for therapy?

1 2 3 4 5 6 7
very unsuccessful very successful
Appendix D

Progress/Outcome Scale: Client Form

On a scale from one to seven rate the outcome of therapy (or progress to this time). Please circle the appropriate number.

How successful has therapy been in achieving your goals for therapy?

1 2 3 4 5 6 7
very unsuccessful very successful