Masters Theses

Graduate College

A Component Analysis of Therapeutic Empathy Behaviors

Barbara P. Nash
Western Michigan University

A COMPONENT ANALYSIS OF THERAPEUTIC
EMPATHY BEHAVIORS

by

Barbara P. Nash

A Thesis
Submitted to the
Faculty of The Graduate College
in partial fulfillment
of the
Degree of Master of Arts

Western Michigan University
Kalamazoo, Michigan
December 1977

ACKNOWLEDGMENTS

I would like to thank Paul Yelsma both for his help and corroboration on this project and for allowing me to use his adaptation of Mehrabian's empathy measure. Special thanks go to Sue Struckens and the Three Rivers Youth Services Center for their unstinting and generous efforts in doing the Carkhuff rating. My thanks to Wayne Fuqua and Malcolm Robertson for their helpful suggestions and criticisms. This work was supported by a grant from the Graduate College Research Fund of Western Michigan University which allowed me to pay my subjects and role players for their generous gifts of time and effort. And of course much thanks is due to the staff of Gryphon Place for their support and encouragement of my work and the unstinting gift of their resources.

I would also like to thank my husband, Barry Ross, for his admonition to "enjoy it," and my parents for their unflagging interest.

Barbara P. Nash

ii

Light flows our war of mocking words, and yet,

Behold, with tears mine eyes are wet!

I feel a nameless sadness o'er me roll.

Yes, yes, we know that we can smile!

But there's a something in this breast,

To which thy light words bring no rest,

And thy gay smiles no anodyne.

Give me thy hand, and hush awhile,

And turn those limpid eyes on mine,

And let me read there, love! thy inmost soul.

Alas! is even love too weak

To unlock the heart, and let it speak?

Are even lovers powerless to reveal

To one another what indeed they feel?

I knew the mass of men concealed

Their thoughts, for fear that if revealed

They would by other men be met

With blank indifference, or with blame reproved:

I knew they lived and moved

Trick'd in disguises, alien to the rest

Of men, and alien to themselves--and yet

The same heart beats in every breast.

                                   -Matthew Arnold

iii

INFORMATION TO USERS

This material was produced from a microfilm copy of the original document. While the most advanced technological means to photograph and reproduce this document have been used, the quality is heavily dependent upon the quality of the original submitted.

The following explanation of techniques is provided to help you understand markings or patterns which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting thru an image and duplicating adjacent pages to insure you complete continuity.

2. When an image on the film is obliterated with a large round black mark, it is an indication that the photographer suspected that the copy may have moved during exposure and thus cause a blurred image. You will find a good image of the page in the adjacent frame.

3. When a map, drawing or chart, etc., was part of the material being photographed the photographer followed a definite method in "sectioning" the material. It is customary to begin photoing at the upper left hand corner of a large sheet and to continue photoing from left to right in equal sections with a small overlap. If necessary, sectioning is continued again — beginning below the first row and continuing on until complete.

4. The majority of users indicate that the textual content is of greatest value, however, a somewhat higher quality reproduction could be made from "photographs" if essential to the understanding of the dissertation. Silver prints of "photographs" may be ordered at additional charge by writing the Order Department, giving the catalog number, title, author and specific pages you wish reproduced.

5. PLEASE NOTE: Some pages may have indistinct print. Filmed as received.

MASTERS THESIS   13-10,939

NASH, Barbara Pritchard
   A COMPONENT ANALYSIS OF THERAPEUTIC
   EMPATHY BEHAVIORS

   Western Michigan University, M.A., 1977
   Psychology, clinical

University Microfilms International, Ann Arbor, Michigan 48106

TABLE OF CONTENTS

iv

CHAPTER I

INTRODUCTION AND REVIEW OF LITERATURE

Empathy, a characteristic of therapist behavior, has been

defined variously: in general terms such as "social sensitivity"

(Bender & Hastorf, 1953); through moderate levels of complexity,

the best example of which is found in Dymond (1949), "the imagina-

tive transposing of oneself into the thinking, feeling and acting of

another and so structuring the world as he does" (for others see

Fromme, Whisenant & Susky, 1974; Hogan, 1969; Mehrabian &

Epstein, 1972; Truax & Carkhuff, 1967); to the extensive definition

given by Rogers (1957). This last definition says, in part, that

empathy involves understanding the client's world as if it were

one's own, moving about in it freely and communicating appropri-

ately this understanding.

It has been so extensively defined because of its hypothesized

importance both as a basic social skill and as an essential thera-

peutic condition. In the former regard it is seen as synonymous

with interpersonal adequacy (Greif & Hogan, 1973) and with inter-

personal effectiveness (Pierce & Zarle, 1972). Dymond (1949)

believes it is basic to an understanding of ourselves and others,

and Hogan (1969) takes it even farther and says that it is essential

1

to the understanding of all social phenomena. Cline (1955) says it is especially crucial in making the judgments about people that one needs to do in business and everyday life.

However, it is the therapeutic value of empathy, as opposed to the social importance, that has been most widely discussed. Although there is some dispute about whether empathy is, in fact, a characteristic of the therapist's behavior (Beutler, Johnson, Neville & Workman, 1972, 1973; Chinsky & Rappaport, 1970), most researchers seem content to proceed as if it is. There have been a small number of studies to substantiate that position (Budman, 1972; Truax, 1961a). These studies suggest that empathy may be impervious to situational variables, and may be a component of a therapist's style of speech.

There has also been disagreement about the construct validity of empathy (Heilman, 1972), but the weight of evidence seems to support the idea that lay people as well as experienced therapists respond consistently to the term and its behavioral manifestation as though empathy were a unified, discrete phenomenon (Hogan, 1969; Shapiro, 1968). As Chandler (1971) says, empathy is a "reliable operationalization of a group of variables which are associated . . . with 'good therapists'."

There have been three basic theories concerning the manner in which empathy might affect therapeutic improvement in the

client. The old psychoanalytic school, exemplified by Katz (1963), believes that empathy enhances diagnostic competence and therapeutic effectiveness. They also believe that it is essential to the development of the therapeutic relationship, upon which the success of therapy ultimately depends. The more recent humanistic school speculates that the essential process leading to therapeutic improvement is that of self-exploration (Rogers & Truax, 1967; Truax, 1961a), and the theory goes, since empathy is one of the core conditions that facilitates self-exploration, it is crucial to therapy (Rogers, 1957; Truax, 1961a; Truax & Carkhuff, 1967). In fact, there is even some rather good experimental evidence to support these ideas (Truax & Carkhuff, 1965). Carkhuff and Berenson (1967) suggest that through empathy the client understands and accepts himself more. Thus empathic responding helps clear up misconceptions that the client might have about himself.

Then there is the behavioristic explanation. There is clear evidence that empathy can be used as a potent reinforcer which clearly increases the frequency of behaviors upon which it is contingent (Truax, 1966b). So empathy can be used selectively to reinforce desired behaviors, or as a modeled behavior to ensure more appropriate social behaviors on the part of the client (Truax, 1966b). Truax and Mitchell (1971) suggested a more extensive behavioral rationale for the use of empathy. They suggested that

empathy could be used (a) to change a client's self-reinforcement schedule, (b) to reinforce self-exploratory behavior so that hidden material would emerge and be amenable to modification, (c) to extinguish anxiety and fear responses to discussions or relationships, and (d) to reinforce human relating and extinguish fear of it.

Regardless of the theoretical foundations of the therapeutic value of empathic therapist behavior, there is extensive evidence that the therapist's use of high levels of empathy is related to client improvement (Betz, 1967; Beutler et al., 1972; Carkhuff, 1969; Carkhuff & Berenson, 1967; Kiesler, Mathieu & Klein, 1967; Truax, 1961a, 1961b, 1963, 1966a, 1966b; Truax & Carkhuff, 1963, 1964, 1965, 1967). And there are few studies to the contrary (Gladstein, 1970; Lesser, 1961). Studies have been done with a wide variety of clients: out-patient neurotics (Truax & Mitchell, 1971), college students (Dickenson & Truax, 1966), elementary school children with behavior problems (Stoffer, 1968), juvenile delinquents (Truax, Wargo & Silber, 1966), and even severe, chronic, hospitalized schizophrenics (Rogers & Truax, 1967). The effect has been shown not only in individual therapy (Truax & Carkhuff, 1967) but also in a group therapy setting (Dickenson & Truax, 1966; Truax, Carkhuff & Kodman, 1965; Truax, Wargo & Silber, 1966). And beneficial results have been obtained even when lay people rather than experienced therapists were trained in

empathy skills (Carkhuff & Truax, 1965b). In fact, Rogers (1957) hypothesizes that empathy produces beneficial effects in all clients, with all therapists, in all settings, and even in life in general. While not going quite that far, Shapiro (1969) concludes that there is a great deal of good correlational data on the subject using varied and sound measures of success. Still, strict experimental evidence is sparse; only Truax and Carkhuff (1965) showed adequate data of this sort. They did an A-B-A design with three subjects and showed that client self-exploration varied a great deal depending upon the levels of empathy emitted by the therapist. This addresses the issue of therapist effectiveness insofar as the goal of this kind of therapy is self-exploration.

If one assumes for the moment that empathy is an important element in therapy, the next question becomes: Can it be effectively taught? There is evidence that people can be taught in relatively short periods of time to function at levels commensurate with those of experienced, empathic therapists (Carkhuff & Truax, 1965b; Truax & Carkhuff, 1965a, 1965b, 1967). This has been done with graduate students and lay people (Carkhuff & Truax, 1965a, 1965b; Shapiro, 1969) as well as with practicing therapists (Truax & Lister, 1971). The techniques widely found to be the most effective in this kind of training have been behavioral ones such as pro-grammed instruction (Magnus, 1973; Saltmarsh, 1973), use of

feedback, modeling and reinforcement (Carlson, 1971, 1974; Dalton, 1973; Fromme et al., 1974; Yoakley, 1972). But there have been more traditional methods that produced good results, too (Carkhuff, 1969; Stern, 1972; Truax, Carkhuff & Douds, 1964). The effects of these training procedures on the various client populations have been quite promising. Carkhuff and Truax (1965a, 1965b) have shown that even inexperienced lay people with no more than a high school education could be trained in empathy skills (in 100 hours) to such an extent that they were able to produce significant improvement in chronic, hospitalized schizophrenics. And, lastly, Carothers and Inslee (1974) have shown that telephone crisis volunteers can be brought to levels of empathic responding commensurate with those of other counselors.

Before being completely confident in these data which show that empathy is an important and learnable skill, one might well ask how client improvement and empathic ability have been assessed. Unfortunately, client improvement has generally been measured in quite unsatisfactory ways. There have been no control groups, behavioral validations or follow-ups to support the personality and projective tests that are used. The MMPI is frequently used (Carkhuff & Truax, 1965a; Kiesler et al., 1967; Truax, Carkhuff & Kodman, 1965) among other inventory-type tests. Other frequently used measures are hospital release and subjective ratings

by nurses and aides (Carkhuff & Truax, 1965a; Kiesler et al., 1967;
Truax, Wargo & Silber, 1966). With out-patient and school children
populations, self-rating (Lesser, 1961) and gains in intelligence and
motivation (Stoffer, 1968) have been used. A critical analysis of
the problems with these measures is beyond the scope of this review,
but there are some obvious faults that may be mentioned. Self-
reports, whether of the personality test sort or other kinds, suffer
from the lack of correlation between verbalizations and other behav-
iors (Risley & Hart, 1968; Rogers-Warren & Baer, 1976). Hospi-
tal release is often more a function of administrative policy than
treatment programs or client improvement (Winett & Winkler,
1972). And subjective rating by others or gains in behaviors
obscurely related to empathic therapy are suspect on validity
grounds. Some of these problems might be overcome or at least
mitigated by the use of control groups, but seldom were they ever
used, not to mention used appropriately (Truax, Wargo & Silber,
1966). I came across no study of this sort that used a systematic
follow-up, and only one that even used the term "behavioral assess-
ment" and this one turned out on closer scrutiny to be a subjective
rating scale with items such as "degree of constructive intraper-
sonal concern" (Carkhuff & Truax, 1965a).

More to the point of this present study is the dearth of good,
objective measures of empathic skills themselves. There are four

basic procedures that have been used in measuring empathy:
(a) accuracy of predictions, (b) personality tests (self-reports),
(c) judged ratings, and (d) behavioral measures. Essentially, the
prediction method involves having the subject observe a client's
behavior and then predict, on that basis, how the person would
respond or feel in various situations. The smaller the discrepancy
between the prediction and the reality, the more empathic the sub-
ject is judged to be (Bender & Hastorf, 1953; Cline, 1955; Danish
& Kagan, 1971; Dymond, 1949; Katz, 1963). Others have used
personality tests to assess empathic ability (Greif & Hogan, 1973;
Hogan, 1969; Lesser, 1961). This involves a person answering
"true" or "false" to self-statements such as "I enjoy the company
of strong-willed people." Then there are the rating scales where
some sort of qualitative criterion of empathy is described and a
judge rates how closely a subject approximates a criterion implicit
in training or explicit in the description. Stern (1972) used such a
measure, as did Guerney, Stover and Demeritt (1968). But by far
the most widely used rating scale is that designed by Truax and
Carkhuff (1967). It was employed by Belluci (1972), Chandler
(1971), Heilman (1972), Shapiro (1968), and Stoffer (1968), to name
a few, not to mention all the studies done by Truax and Carkhuff
themselves (see references). (It may be obvious that not only do
these measures differ widely in their approach, but there may even

be serious problems in the equation of these three types of measures since they all involve different behaviors with different discriminative stimuli and reinforcers. Nonetheless, they are all used with varying frequency.)

The Carkhuff-Truax scale itself consists of qualitative descriptions of various levels of empathic responding from the lowest level, "feeling and meaning both absent or inaccurate," to the highest, "personalized goal--response captures goal which is flip side of personalized problem; also recaptures personalized problem." Its use involves the employment of two or three judges to rate segments of transcripts from therapy sessions. They give each segment a number, and the raters' numbers are averaged for a final figure.

The problems with these kinds of measures bears some mention. Predictive tests, though straightforward, are fraught with problems. Bender and Hastorf (1953) point out that the more similar people are, the more their scores on these measures tend to be spuriously inflated due to a general tendency of subjects to predict that others will respond as they do. These measures in general are subjective and cumbersome to use. They often rely on vague mentalistic concepts which are seldom operationally defined, and thus suffer from semantic problems. Hogan (1969) cites other major methodological and validity problems in one of the better predictive empathy measures (i.e., Dymond's).

The measures that follow the personality-test format (e.g., Barrett-Lennard, 1962) have other characteristic difficulties. We are interested in the behavior of being empathic, and these tests measure only one's verbal responses about behavior. Since there is no evidence that verbal responses correlate highly with actual behavior (Risley & Hart, 1968; Rogers-Warren & Baer, 1976), we cannot rely on these measures too heavily. More specifically, the MMPI does not even have a specific empathy scale on it, and it was never validated for this purpose. Danish and Kagan (1971) studied one such test, administering it before and after an empathy training weekend, and found slight differences in the means, but wide individual variation (sometimes in a direction opposite that predicted).

The judged rating scales--the Carkhuff-Truax one in particular--have been criticized on a number of grounds. Beutler et al. (1972) say they are still far too subjective. The validity of the scale itself has never been examined; validity has only been measured via the effect on clients of those who score highly on the test (Shapiro, 1969). In fact, it remains to be shown that the scale does not reflect qualities of the therapist-client dyad rather than purely those of the therapist; there are data to the contrary (Beutler et al., 1973). If the scale does reflect qualities of the dyad, the interrater reliabilities (Truax, 1972) are calculated in a way that could lead to spuriously high results (Beutler et al., 1972; Chinsky & Rappaport,

1970). And the rate-rerate reliabilities are unacceptably low: .70 (Carkhuff & Truax, 1965b).

There is a great deal of speculation that the ratings reflect many things other than empathy itself; that is, there are validity problems in the measure itself. (This may relate to poorly operationalized definitions.) Some say it reflects the therapist's voice qualities such as inflection, language style, tone of voice, phraseology, etc. (Chinsky & Rappaport, 1970; Shapiro, 1969). Others say it reflects the therapist's commitment to or intensity of involvement with the client, or the amount of client verbalization (Kiesler et al., 1967), but that in any event the results could be confounded by or even an artifact of the rating procedure. The fact that deleting the client's half of the dialogue in the rated transcripts had little or no effect upon the ratings given makes this scale somewhat suspect (Truax, 1966). It suggests that the raters are responding to the form or stereotype of therapist response and ignoring such things as the importance of timing, etc.

In short, these three most used forms of empathy measures all suffer from varying degrees of debilitating flaws. They are subjective, vague, and inaccurate. Partly for these reasons their inventors and users needed to resort to statistical manipulations, factor analyses, etc., to demonstrate their value. Michael (1974) points out the manifold problems with, and faults in the use of

statistics in demonstrating treatment effects. But even granting the limited usefulness of these tools, the statistics themselves, when used, do not show a significant correlation among the major empathy tests (Heilman, 1972; Resnikoff, 1972). And, as Eisler (1976) says, it is hard even to achieve high interrater reliability with such subjective measures.

To remedy this situation one might well look to behavioral methods of assessment. There are surprisingly few mentions of this possibility. Pierce and Zarle (1972) looked at one factor: How many times a subject mentioned the feelings of others during a session. Hargrove (1974) looked at four factors: (a) duration of utterance, (b) reaction time latency, (c) initiative time latency, and (d) interruptions. And Guerney et al. (1968) looked at the proportion of reflective statements out of the total number of statements made during a session. (They all used some other types of measures to show that their subjects were empathic.) These results taken together are suggestive that empathic people might attend more to the feelings of others, and reflect them in their speech, interrupt less, and allow more silence. But these studies individually are quite weak and merely suggestive. Pierce and Zarle (1972) and Guerney et al. (1968) each looked only at one factor. And Hargrove (1974) tells us mainly what empathic therapists do not do. As he puts it, "silence behaviors are . . . indicative of

high levels of empathy." So there still remains, it seems to me, a great deal of research to be done on the behavioral measurement of empathic skills.

My purpose in this experiment was to replicate these earlier findings, and to discover new behaviors that might be related to empathy skills, and might have functional relationships with each other. For the following reasons, I also evaluated the subjects with the Carkhuff-Truax scale and a personality test of empathy. The first reason involves external validity: I wanted to be sure that according to the widely used definition, empathy was, in fact, being learned by the subjects. The other reason is best described by Eisler (1976):

> In most assessment situations requiring behavioral meas-
> ures of social skill, it is probably better to use a combi-
> nation of objective and subjective measures rather than
> relying on either type alone. While it is generally more
> difficult to obtain high interjudge agreement with sub-
> jective measures they appear to possess greater social
> validity than any single objective behavioral measure of
> social skill. This is probably due to the fact that raters
> respond to subtle behavioral cues and social contextual
> variables which usually cannot be specified with objective
> measures. (p. 374)

As previously stated, the purpose of this experiment was to examine some specific behaviors to see whether they changed in frequency as a person went through empathy training, and thus whether these might be some of the behaviors a person emits when s/he is being empathic. Understanding clearly the component behaviors of empathic responding would better enable us to isolate the effective components, to train people in this skill, and to develop objective measures of empathy thus allowing quantification of the relationship between client progress and the complex set of behaviors called empathy. The behaviors I chose were five in number. The first was simply the percentage of time (per session) the therapist (subject) spent talking. The therapist's responses were defined as any unbroken statement (sentence or phrase) that conveyed one thought. The second and third factors were thus derived from the total number of therapist responses defined this way. The second factor was the percentage of reflective statements made by the therapist. The third factor was the percentage of questions asked by the therapist. The fourth behavior was the number of client agreements or clarifications (e.g., "well, that's not exactly it, it's more like . . .") per therapist response. The last category was simply the number of feeling words (or emotion words) used by the therapist per five minute interval. (The form is shown in Figure 1.)

The literature and my own conjectures suggested what to

Figure 1

Behavioral Measure Form

DURATION OF SUBJECT'S
SPEECHES

TOTAL LENGTH OF
ROLE PLAY

NUMBER OF REFLECTIVE
SUBJECT'S RESPONSES
(you think, you feel, . . .)

QUESTIONS

OTHER RESPONSES

NUMBER OF SPECIFIC
FEELING WORDS
(used by subject)

NUMBER OF AGREEMENTS
OR CLARIFICATIONS
(by confederate)

expect in the trends of these behaviors. The work of Hargrove (1974) suggests that empathy involves a great deal of listening and less talking on the part of therapists, so I guessed that the percentage of time they spent talking would decrease during training. Guerney et al. (1968) suggest that empathic behavior would involve the therapists making more reflective responses (such as "You were really upset by that"). Pierce and Zarle's (1972) work led me to predict that the therapists would use more specific feeling words ("good" and "bad" did not count) as they learned to be more empathic. I predicted that they would ask fewer questions; and as they refined their empathy skills, they would elicit more agreements and/or clarifications from the clients. (I also predicted that these gains would all decline one month after training ended since the prompts would no longer be present.)

My rationales may be fairly obvious, but I will describe them briefly. In order to gain a clear understanding of a client's situation and feelings, therapists must do a great deal of listening. In order to be sure they are understanding correctly and in order to communicate that effort (and its relative success or failure), reflecting back what is understood is essential. This process also tends to elicit more information (in a non-threatening way) from the client in the form of agreements and/or clarifications, which obviates the necessity of the therapists asking many direct (and perhaps

threatening) questions. And in the process of reflecting back and clarifying what the client feels about his or her situation, the therapists may well use more specific emotional words than they would in other forms of therapy. (The factor of agreements or clarifications from the clients was also included so that not only what Mehrabian and Epstein (1972) refer to as the transmission of empathy, but also its reception would be involved in this empathy measure.) These, then, were my expectations as I entered the experiment. The expectation about a decrease in gains is based on the fact that after training ends there are no more structured prompts or positive reinforcers to maintain the behaviors. In that situation one might expect that the behaviors would tend to follow an extinction curve (Ferster & Skinner, 1957).

CHAPTER II

METHOD

Subjects

This experiment was conducted at Gryphon Place, a crisis
intervention center in Kalamazoo, Michigan. The subjects were
experienced volunteers and trainees at the center. They were all
in their 20's. Males and females participated in fairly equal num-
bers. Their professional backgrounds ranged from undergraduate
students to school psychologists and doctoral candidates, all tending
to be in the helping professions.

Apparatus

Three instruments were used to evaluate the empathy skills of
the subjects. The Carkhuff-Truax Accurate Empathy Scale (Truax
& Carkhuff, 1967) was scored by judges trained in its usage. Also
a paper and pencil empathy measure was used. This test was
adapted by Yelsma[1] from an empathy measure developed by
Mehrabian and Epstein (1972).[2] (For brevity's sake, this will be
referred to as the Yelsma Test.) The third measurement instru-
ment was a behavioral checklist. (See Figure 1.) Duration was
measured by a stopwatch and included any verbalization made by the

18

subject. Total length of role play was also determined by stop-watch. Reflective responses were defined as a response that contained information which the subject had already given. A question was any response that asked for information not already given. Other responses were any other phrase or sentence (not "uh huh," or "really?"). Specific feeling words (that is, not "good" or "bad") were counted only if they referred to the client or others involved in the problem situation. Agreements or clarifications were scored if the clients responded positively to a therapist statement, or if they responded by giving more information on the subject at hand.

The stimulus used was a role play of a crisis call. Five role-players were each trained in one role until they could perform it consistently. The roles were one-paragraph descriptions of a typical crisis-type problem. (The ideas came from Katz, 1975.[3]) The role plays were conducted over an intercom telephone line in the center itself (participants were in separate rooms). They were recorded on a cassette tape recorder from a small, suction-cup microphone. The tapes were rated on the checklist shown in Figure 1. A reliability rater was trained to 80 percent agreement and rated one-third of the tapes.

The empathy training itself was done by the Gryphon Place staff. It lasted 35 hours over a 2-week period. The specific skills taught included practicing identifying feelings and content in a

person's problem, reflecting feelings and content, clarifying what one does and does not understand and getting needed information in a non-threatening way, identifying goals and wishes, learning to avoid "blocks" in communications and some specific problem-solving strategies. Each session would begin with a lecture and demonstration, followed by small-group exercises to practice the particular skills. (These groups were led and structured by two trained leaders, and had four to five trainees.) The practices would involve role plays and feedback from all members (including trainers) of the group. Large group discussions would follow, and end each training session.

Procedure

The dependent variables were as mentioned above: (a) duration of therapist's speaking, (b) percentage of therapist's reflective statements, (c) percentage of therapist's questions, (d) agreements or clarification elicited from the client, and (e) number of specific feeling words used by therapist. The independent variable was the empathy training by the Gryphon Place staff. The experimental design was an A-B-A design; that is, before training, during and immediately after training, and one month after training (with a comparison group used). The three measures of empathic skill were administered as in the following chart and explanation.

Time	I Before Training	II After Phase 1	III After Phase 2	IV After Phase 3	V One month Follow-up
Test Adminis- tered	Behavior Measure Yelsma Test Carkhuff Test	Behavior Measure	Behavior Measure	Behavior Measure Yelsma Test Carkhuff Test	Behavior Measure

The subjects were divided into two groups: experimental and comparison. The comparison group consisted of 11 females and 10 males who were chosen by the director, from the ranks of experienced crisis workers at the center, for their mastery of empathy skills. They had all been through the training previously. The experimental group comprised 9 people, 6 male and 3 female, who were chosen at random from the group of trainees at the center. There were also five experienced volunteers chosen and trained to role-play the crisis call.

The three measures were administered as follows:

1. Yelsma Test. This paper and pencil test was given once to the comparison group and twice to the experimental subjects, before and immediately after training.

2. Behavioral Measure. This measure was used to record

the previously described behaviors during five role plays of the crisis calls. The role plays were done before training, one after each of the three phases of training, and one a month later. The instructions given to the subjects were simply to treat the call as if they were already a volunteer on the crisis line. Each subject had all five role plays in a randomized order. The tapes were rated blindly so that the raters did not know whether a tape came from before training, during it, or afterwards. The reliability rater was trained to a criterion of 80 percent agreement using Type I reliability. This method of calculating reliability involves summing the frequencies of each category of behavior and dividing the smaller number into the larger number for each category. The rater did one-third of the tapes. Her training consisted of a number of phases. The first was observing my rating of two tapes, and hearing the verbal descriptions of the response categories. (This was not sufficient to reach 80 percent agreement.) Then she rated tapes with my comments occurring only when she committed an error. Lastly, she rated practice tapes by herself. When our agreement reached 80 percent, the reliabilities rating began.

3. <u>Carkhuff Rating</u>. The first and fourth tapes of the role plays were also rated by trained Carkhuff judges. They listened to the first five or ten minutes of each tape and rated each therapist response according to the descriptions shown in Figure 2. The

Figure 2

Carkhuff-Truax Accurate Empathy Measure

Level System

Level 1.0 - Feeling and meaning both absent or inaccurate.

Level 1.5 - Accurate response to content, feeling absent or inaccurate.

Level 2.0 - Accurate response to meaning, feeling absent or inaccurate.

Level 2.5 - Accurate response to feeling, meaning absent or inaccurate.

Level 3.0 - Accurate response to feeling, accurate response to meaning.

Level 3.5 - Personalized meaning--response captures the role the helpee is playing in problem.

Level 4.0 - Personalized problem--response captures the behavioral defecit of helpee which is causing the problem; also captures feeling helpee has of self as result of defecit.

Level 4.5 - Personalized goal--response captures goal which is "flip side" of personalized problem; also recaptures personalized problem.

Note. From The Art of Helping Trainer's Guide by R. R. Carkhuff & R. Pierce. Human Resource Development Press, 1975.

response levels in each segment (one per tape) were then averaged

resulting in a final overall figure for that tape. These ratings

were also done blindly so that the raters did not know which tape

was first and which was fourth.

CHAPTER III

RESULTS

Behavioral Measure

The group averages here showed that (a) the amount of time the subjects spent talking declined, (b) the percentage of reflective statements increased, (c) the percentage of questions asked did not change, (d) the subjects elicited more agreements or clarifications from their "clients," and (e) the subjects used fewer feeling words overall. These averages come from comparing before training to after training (I - IV). (See Figure 3.)

The absolute numbers of subjects who followed these trends are consistent with these averages: Of the nine experimental subjects, seven spoke less, six used fewer feeling words, six used more reflective statements, and six got more agreements from their clients. (See Table 1.) Those subjects whose behavior opposed these trends tended to change very little, relative to the others' changes. In all analyses, the number of questions asked changed little and showed no clear trends. For that reason I exclude it from further comments.

The overall results of my follow-up one month later were

25

Figure 3

26

Group Averages on Behavioral Measure
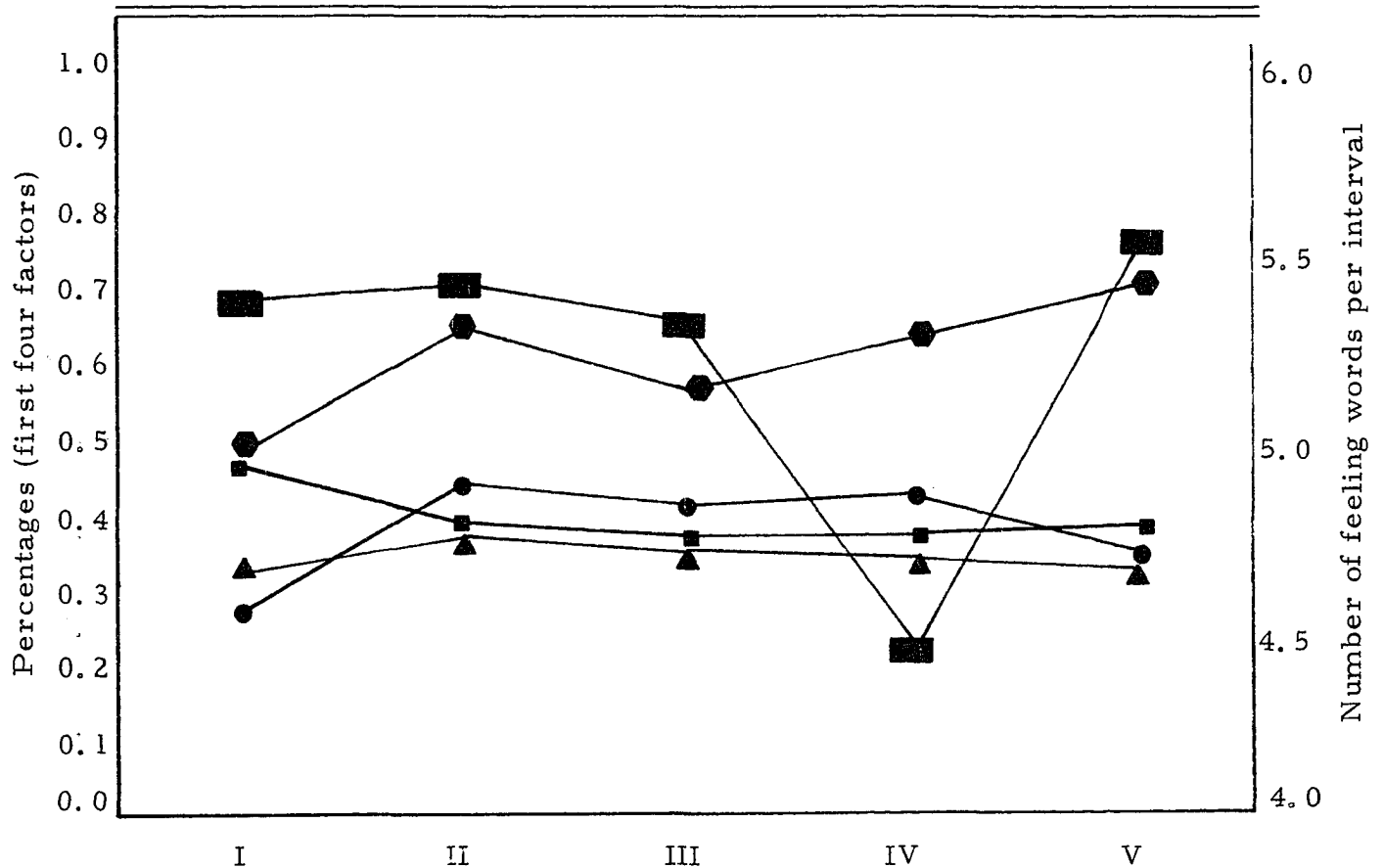
Averages for nine experimental subjects on behavioral measure before training, over the three phases of training, and one month later. First four factors relate to percent axis, last factor, to right axis.

■   Percentage of time therapist spoke

●   Percent reflective statements

▲   Percent questions

⬢   Percent client agreement per therapist statement

▦   Number of feeling words used by therapist per interval

Table 1

Changes in Frequencies of Behaviors (I to IV)

Scores Rose on Y-test	Duration			Reflectives			Questions			Agreements			Feeling Words		
	I	IV	Change	I	IV	Change	I	IV	Change	I	IV	Change	I	IV	Change
1	.43	.40	-.03	.17	.42	+.25	.45	.35	-.10	.29	.62	+.33	3.33	7.00	+3.67
2	.58	.42	-.16	.04	.42	+.38	.27	.51	+.24	.23	.64	+.41	3.00	5.50	+2.50
5	.38	.39	+.01	.38	.38	.00	.27	.34	+.07	.46	.55	+.09	5.67	6.33	+ .66
7	.29	.27	-.02	.09	.43	+.34	.42	.33	-.09	.81	.67	-.14	6.00	2.50	-3.50
8	.63	.25	-.38	.15	.52	+.37	.22	.45	+.23	.28	.83	+.55	5.60	2.67	-2.93
Scores Fell on Y-test															
3	.38	.25	-.13	.26	.34	+.08	.26	.23	-.03	.37	.69	+.32	6.00	3.50	-2.50
4	.41	.44	+.03	.64	.48	-.16	.06	.27	+.21	.61	.58	-.03	8.50	6.50	-2.00
6	.50	.47	-.03	.43	.37	-.06	.34	.18	-.16	.59	.37	-.22	5.50	4.33	-1.17
9	.45	.26	-.19	.23	.38	+.15	.46	.29	-.17	.57	.62	+.05	4.50	1.50	-3.00

Note. First four factors are percentages; the last factor is number per interval.

equivocal. On all factors, five subjects increased their behaviors and four decreased. (See Table 2.) The averages showed that the trends in duration and reflective statements reversed. On the other hand, the agreements from clients increased as did the number of feeling words used (the latter increased dramatically). (See Figure 3.)

Looking at individual subjects' behavior, there were two subjects (1 and 2) who performed as the literature and my own reasoning would lead us to expect. That is, they spoke less, reflected more, got more agreements, and used more feeling words after training than before it. (These two subjects are discussed again below.) The absolute highest scores, however, went to subject 8 in three of the four categories. (Subject 1 used the most feeling words.) (This is significant because the same subject also scored highest on the other tests--see below.)

An interesting fact becomes apparent when we look at the subject who, in terms of trends, performed against expectations in all categories. Subject 4 spoke more, reflected less, got fewer agreements, and used fewer feeling words after training than before it; but she started (before training) with extremely high ratings on all these factors. She used by far the most reflective statements and feeling words of any subject before training, got a great number of agreements and spoke very little. (See Table 1.) Subject 6 also

Table 2

Changes in Frequencies of Behaviors (IV to V)

Scores Rose on Y-test	Duration			Reflectives			Questions			Agreements			Feeling Words		
	IV	V	Change	IV	V	Change	IV	V	Change	IV	V	Change	IV	V	Change
1	.40	.30	-.10	.42	.35	-.07	.35	.41	+.06	.62	.71	+.09	7.00	6.00	-.100
2	.42	.44	+.02	.42	.30	-.12	.51	.30	-.21	.64	.56	-.08	5.50	4.33	-1.17
5	.39	.30	-.09	.38	.40	+.02	.34	.20	-.14	.55	.72	+.17	6.33	3.00	-3.33
7	.27	.46	+.19	.43	.13	-.30	.33	.49	+.16	.67	.49	-.18	2.50	1.67	- .83
8	.25	.38	+.13	.52	.20	-.32	.45	.20	-.25	.83	.60	-.23	2.67	5.00	+2.33
Scores Fell on Y-test															
3	.25	.44	+.19	.34	.40	+.06	.23	.33	+.10	.69	.59	-.10	3.50	4.33	+ .83
4	.44	.30	-.14	.48	.81	+.33	.27	.15	-.12	.58	.88	+.30	6.50	9.00	+2.50
6	.47	.34	-.13	.37	.41	+.04	.18	.47	+.29	.37	.82	+.45	4.33	6.50	+2.17
9	.26	.38	+.12	.38	.46	+.08	.29	.33	+.04	.62	.73	+.11	1.50	9.67	+8.17

performed against most expectations, and he will be discussed below also.

Unfortunately, the two subjects who performed as expected in every category on this scale (1 and 2) both showed the most reversals at follow-up. On reflective statements and feeling words, both reversed, and one each reversed on the two other factors.

Reliability

My interrater reliability figures for the behavioral ratings averaged to .81 using Type I reliability (Bijou, Peterson & Ault, 1968). The breakdown is as follows: duration .87, reflectives .77, questions .86, agreements .87, and feeling words .68.

Self-Report (Yelsma Test)

When we analyze these results in terms of Yelsma's personality test, we find that it separates the subjects into two distinct groups in terms of their behaviors. (For scores, see Table 3.) From before training to afterward, the two people whose scores increased the most on this empathy measure were those who performed according to expectations on the behavioral scale (1 and 2). The highest overall scorer was also the highest scorer on my scale (Subject 8). Some subjects increased their scores on Yelsma's test of empathy (five out of nine), but there was not an

Table 3

Scores on Yelsma Test

Subjects	Pre-Training, $\overline{X}$ = 43.5	Post-Training, $\overline{X}$ = 39.3
1	24	37
2	22	36
3	54	22
4	58	52
5	51	58
6	27	9
7	30	33
8	63	73
9	63	34
Comparison Group $\overline{X}$	48.67	
Range	27 - 79	

unequivocal trend. The comparison group scored a good deal higher than these trainees whether we look at before or after train-ing. (See Table 3.) But when we look at the differences between those whose scores rose vs. those whose declined, we do see some definite trends. (See Tables 1 and 2.) Those whose scores increased tended to speak less (four out of four), elicit more agreements (four of five), and use more feeling words (three of five). (In the last two categories, Subject 7 was the contrary one, and her score started very low and increased very little.) Those whose scores declined talked more (three of four) and used fewer feeling words (four of four) after training than before it. Looking at the follow-up data shows us that, unfortunately, the subjects who performed as expected tended to reverse their behaviors more than the others. On the other hand, the others tended to improve their behavior over time. (See Table 2.)

Yelsma's test correlated moderately well with three of four behaviors: duration -.52, reflectives .69, agreements .51, and feeling words .26.

So the interaction of these two scales tells us more about what to expect from the subjects than either test alone. Of the three people who used more feeling words after training, all of them had increased empathy scores on Yelsma's test. The two whose elicited agreements declined a great deal had low scores

after training. And Subject 4, whose behaviors <u>before</u> training were what we expected from subjects <u>after</u> empathy training, and who tended to perform against expectations later on, also scored very highly on this test and dropped her score after training. Subject 6, who was the other subject that performed the most against expectations on the behavioral measure, scored quite low on this test and his score dropped abysmally after training.

There were two subjects whose scores were not consonant with their behaviors. Subject 7, whose score on this test rose, did not perform as expected on the behavioral measure, but her score started quite low and went up only slightly. And there was Subject 3, whose scores fell on this test; but on the other measure she performed much like Subjects 1 and 2 (except for feeling words). Her score on the Yelsma Test began quite high.

Carkhuff Ratings

The Carkhuff Test also adds some information to these results. (See Table 4.) Subject 1 who scored as expected on the behavioral measure, gained the most on this test. Subject 8, the high scorer on both other tests, had the highest score on this test also. Overall, all but one subject gained in rated empathy on this test from before training to after it. As one might expect, this one was Subject 4, who started with a higher score on this test than all but

Table 4

Scores on Carkhuff Measure

Subjects	Pre-Training	Post-Training
1	1.0	2.0
2	1.0	1.5
3	1.5	2.0
4	2.5	2.0
5	1.0	1.5
6	1.0	1.5
7	1.0	1.5
8	1.5	2.5
9	1.0	1.5

one of the subjects had even <u>after</u> training. Most subjects gained

very little (.5) on this test. In fact this is the smallest increase

possible on this scale. Subjects 1 and 8 gained 1.0. (These were

ones that performed as expected on the other tests, also.)

The Carkhuff Test correlated best with feeling words on the

behavioral measure (.57). It correlated less well with other

behaviors: duration -.01, reflectives .49, and agreements .37.

These last two measures (Yelsma and Carkhuff) correlated with

each other only .28. (The three strongest factors on the behavioral

test correlated well with each other: duration with reflectives -.59,

reflectives with agreements .57, and duration with agreements

-69.)

CHAPTER IV

DISCUSSION

My data show that as subjects go through an empathy training and as their behaviors become more empathic according to a widely used criterion, certain specific behaviors change as the literature and my own reasoning suggest they would. These people tend to speak less, make more reflective statements, and elicit more agreements and/or clarifications from their clients. There was also some evidence (in terms of correlations) that these people use more feeling words, also. The two most frequently used measures of empathy do correlate with these behaviors, albeit weakly in some cases. And each test correlated much more strongly with some behaviors than with others. The question-asking behavior of the subjects did not change appreciably. I suspect (and this may be a suggestion for further research) that these subjects learned to ask for information not less frequently, but in a less direct, less threatening way. That is, one might investigate how the form of questions changes during empathy training. It may be that instead of saying, for example, "What happened then?", they might say "I'm a little unclear about what happened then."

The three forms of empathy measures, though they correlated

36

weakly, did tend to agree on evaluations of certain subjects' behavior. They all pointed toward the same people and the same changes in style. That is, they agreed on the subjects who performed at the highest and lowest levels. I think the Carkhuff scale offered the least information; it correlated less well with the behaviors than the other test, and it showed very few differences between subjects. The fact that they agreed and correlated so well is surprising in light of the fact that they all are really measuring very different behaviors. At present there is no way to know whether the behaviors counted are the same behaviors that lead to high ratings on the Carkhuff scale. And the Yelsma Test clearly measures quite different responses.

My data on reversals were not consistent across subjects. They showed that some people do not tend to maintain the gains made through training (especially those who made the highest gains), but some do. So further research on maintenance of these behaviors might be in order. It might also have been helpful to give Yelsma's test at follow-up in order better to correlate the behavior changes with these scores. It is interesting that the average level of agreements and use of feeling words did tend to increase at follow-up. Perhaps this means that some of the subjects are refining their skills as they gain more experience, or that other variables are controlling their behaviors now, besides the training.

It is also reassuring in that it suggests that empathy is not only being emitted (or sent) by the therapist, but it is also being received by the client.

I would also suggest from the singular behavior of Subject 4 that this training is not useful for people who are highly empathic to begin with. They might be screened beforehand and allowed to forego the training. Otherwise their high level of empathy may tend to decrease (see her behavioral trends and her score on Yelsma's test). I suspect this is because she felt constrained to express her empathy in ways that were not habitual to her. The modeling done during the training focused on certain specific ways of phrasing reflections and questions. This may have made her self-conscious about her usual phraseology.

The fact that Yelsma's test showed the comparison group to be higher than the trainees both before and after training (see Table 3) suggests three possible explanations. The most likely one is that these people were selected by the director as being high on empathy skills; he probably knew what he was doing. There could be a self-selection factor at work here, too; those who stay at the center for any length of time may tend to be the more successful ones. Or perhaps people do refine their skills as they work there, and the differences are a result of their longer experience at the center. It would have been informative to do a behavioral analysis

of these older volunteers' behavior as I did for the trainees.

Yelsma's test gave more information about the behaviors chosen, as it correlated well with three of the four targeted behaviors. The Carkhuff Test focused more on feeling words used and to a lesser extent on reflective statements. The fact that each test correlated better with all but one of my factors than they did with each other suggests that I may have found elements common to both tests. The fact that three of my factors correlated so well among themselves suggests that they may all be part of a larger whole: empathy.

My work does tend to substantiate the work of the behavioral researchers mentioned in the literature review. Pierce and Zarle's (1972) work and mine both show that the subjects do mention others' feelings more as they become more empathic. I found, as did Hargrove (1974) that empathic people speak less. And the work of Guerney et al. (1968) shows, with mine, that these people make a higher proportion of reflective statements. The theoretical contribution made here (besides replication) involves the additional factor of agreements and clarifications, as well as the suggestion that we look at the form of questions asked during a session, and perhaps other specific behaviors. It also suggests that empathy behaviors are amenable to objective study. The practical implications of the work involve screening out high empathy trainees, focusing more

specifically in training on the behaviors that seem to reflect high empathy and creating some safeguards to prevent reversals of training effects. Any of these tests might also be used to do some post-training evaluations as a form of behavioral or non-behavioral criterion for admission to the staff.

These results point to the importance of trying to develop better behavioral definitions of empathy for training purposes and in order to facilitate empirical work on a quantitative level about empathy and its effects on client outcomes. Trainees should be taught to decrease their own speaking behavior ("listen more," or "keep the focus on the client"), to use more reflective statements, to learn new ways of asking for information, to be aware of the client's responses as feedback, and to use more specific feeling words. Other types of empathy measures may be preferred for evaluations due to their convenience, and because these are useful in choosing the most empathic people from a group. But they do not tell us which behaviors are weaker than others, and thus, which ones to emphasize in training. For this same reason they are not as useful in evaluating the specific benefits or weaknesses of a training program; that is, they tell us only global scores, not on which behaviors the trainees are strong or weak.

A bit of anecdotal information about the effects of this study might be in order here. Those trainees who were part of the study

and did the role plays found them to be a useful yardstick for a self-evaluation of their progress. They thought all trainees should do them. Strangely enough, even though the order of the role plays was different for each subject, most commented that the later ones seemed harder. I speculate that this is due to their increased awareness of the complexities of the tasks.

If we are to continue believing in the efficacy of empathy skills, we should continue the endeavor to specify what these skills are and which are most important in improving therapeutic out-comes.

References

Barrett-Lennard, G. T. Dimensions of therapist response as causal factors in therapeutic change. _Psychological Monographs_, 1962, _76_, (43, Whole No. 562).

Bender, I. E., & Hastorf, A. H. On measuring generalized empathic ability (social sensitivity). _Journal of Abnormal and Social Psychology_, 1953, _48_, 503-506.

Belluci, J. E. The development and validation of a measure of counselor empathy. _Dissertation Abstracts International_, 1972, _32_, 4933A-4934A.

Betz, B. J. Studies of the therapist's role in the treatment of the schizophrenic patient. _American Journal of Psychiatry_, 1967, _123_, 963-971.

Beutler, L. E., Johnson, D. T., Neville, C. W., & Workman, S. N. "Accurate empathy" and the A-B dichotomy. _Journal of Consulting and Clinical Psychology_, 1972, _38_, 372-375.

Beutler, L. E., Johnson, D. T., Neville, C. W., & Workman, S. N. Some sources of variance in "accurate empathy" ratings. _Journal of Consulting and Clinical Psychology_, 1973, _40_, 167-169.

Bijou, S. W., Peterson, R. F., & Ault, M. H. A method to integrate descriptive and experimental field studies at the level of

data and empirical concepts. Journal of Applied Behavior Analysis, 1968, 1, 175-191.

Budman, S. H. Client's lexical organization and psychotherapist's empathy level. Psychological Reports, 1972, 31, 77-78.

Carkhuff, R. R. Helping and human relations: A primer for lay and professional helpers. New York: Holt, Rinehart & Winston, 1969.

Carkhuff, R. R., & Berenson, B. G. Beyond counseling and therapy. New York: Holt, Rinehart & Winston, 1967.

Carkhuff, R. R., & Truax, C. B. Lay mental health counseling: The effects of lay group counseling. Journal of Consulting Psychology, 1965, 29, 426-431. (a)

Carkhuff, R. R., & Truax, C. B. Training in counseling and psychotherapy: An evaluation of an integrated didactic and experiential approach. Journal of Consulting Psychology, 1965, 29, 333-336. (b)

Carlson, K. W. Reinforcement of empathy: An operant paradigm for the training of counselors. Dissertation Abstracts International, 1971, 32, 168A-169A.

Carlson, K. W. Increasing verbal empathy as a function of feedback and instruction. Counselor Education and Supervision, 1974, 13, 208-213.

Carothers, J. E., & Inslee, L. J. Levels of empathic understanding offered by volunteer telephone services. Journal of Counseling Psychology, 1974, 21, 274-276.

Chandler, D. R. Construct validity of a scale of accurate empathy in psychotherapists. Dissertation Abstracts International, 1971, 32, 1836B-1837B.

Chinsky, J. M., & Rappaport, J. Brief critique of the meaning and reliability of "accurate empathy" ratings. Psychological Bulletin, 1970, 73, 379-382.

Cline, V. B. Ability to judge personality assessed with a stress interview and sound-film technique. Journal of Abnormal and Social Psychology, 1955, 50, 183-187.

Dalton, R. F. An application of principles of social learning to training in communication of empathy. Dissertation Abstracts International, 1973, 33, 5597A.

Danish, S. J., & Kagan, N. Measurement of affective sensitivity: Toward a valid measure of interpersonal perception. Journal of Counseling Psychology, 1971, 18, 51-54.

Dickenson, W. A., & Truax, C. B. Group counseling with college underachievers. Personnel and Guidance Journal, 1966, 45, 243-247.

Dymond, R. A scale for the measurement of empathic ability. Journal of Consulting Psychology, 1949, 13, 127-133.

Eisler, R. M. Behavioral assessment of social skills. In M.

Hersen & A. S. Bellack (Eds.), Behavioral assessment.

New York: Pergaman Press, 1976.

Ferster, C. B., & Skinner, B. F. Schedules of reinforcement.

New York: Appleton-Century-Crofts, 1957.

Fromme, D. K., Whisenant, W. F., & Susky, H. H. Group

modification of affective verbalizations. Journal of Consulting

and Clinical Psychology, 1974, 42, 866-871.

Gladstein, G. A. Is empathy important in counseling? Personnel

and Guidance Journal, 1970, 48, 823-827.

Greif, E. B., & Hogan, R. The theory and measurement of

empathy. Journal of Counseling Psychology, 1973, 20,

280-284.

Hargrove, D. S. Verbal interaction analysis of empathic and non-

empathic responses of therapists. Journal of Consulting and

Clinical Psychology, 1974, 42, 305.

Heilman, K. A. Empathy: The construct and its measurement.

Dissertation Abstracts International, 1972, 33, 2791B.

Hogan, R. Development of an empathy scale. Journal of Consulting

and Clinical Psychology, 1969, 33, 307-316.

Katz, R. L. Empathy: Its nature and uses. London: Collier-

Macmillan, 1963.

Kiesler, D. J., Mathieu, P. L., & Klein, M. H. A summary of the issues and conclusions. In C. R. Rogers, E. T. Gendlin, D. J. Kiesler & C. B. Truax (Eds.), The therapeutic relationship and its impact: A study of psychotherapy with schizophrenics. Madison, Wisc.: University of Wisconsin Press, 1967.

Lesser, W. M. The relationship between counseling progress and empathic understanding. Journal of Counseling Psychology, 1961, 8, 330-336.

Magnus, R. E. Development of an individualized auto-instructional program to improve the ability to understand and communicate empathically. Dissertation Abstracts International, 1973, 33, 3296A.

Means, B. L. Levels of empathic response. Personnel and Guidance Journal, 1973, 52, 23-28.

Mehrabian, A., & Epstein, N. A measure of emotional empathy. Journal of Personality, 1972, 40, 525-543.

Michael, J. Statistical inference for individual organism research: Mixed blessing or curse? Journal of Applied Behavior Analysis, 1974, 7, 647-653.

Pierce, R. M., & Zarle, T. H. Differential referral to significant others as a function of interpersonal effectiveness. Journal of Clinical Psychology, 1972, 28, 230-232.

Resnikoff, A. A measure of discrimination for the prediction of counselor effects. Journal of Counseling Psychology, 1972, 19, 464-465.

Risley, T. R., & Hart, B. Developing correspondence between the non-verbal and verbal behavior of preschool children. Journal of Applied Behavior Analysis, 1972, 5, 449-504.

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

Rogers, C. R., & Truax, C. B. The therapeutic conditions antecedent to change: A theoretical view. In C. R. Rogers, E. T. Gendlin, D. J. Kiesler & C. B. Truax (Eds.), The therapeutic relationship and its impact: A study of psychotherapy with schizophrenics. Madison, Wisc.: University of Wisconsin Press, 1967.

Rogers-Warren, A., & Baer, D. M. Correspondence between saying and doing: Teaching children to share and praise. Journal of Applied Behavior Analysis, 1976, 9, 335-354.

Saltmarsh, R. E. Development of empathic interview skills through programmed instruction. Journal of Counseling Psychology, 1973, 20, 375-377.

Shapiro, D. A. Empathy, warmth and genuineness in psychotherapy. British Journal of Social and Clinical Psychology, 1969,

<u>8</u>, 350-361.

Shapiro, J. G. Relationship between expert and neophyte ratings of therapeutic conditions. <u>Journal of Consulting and Clinical Psychology</u>, 1968, <u>32</u>, 87-89.

Stern, J. The effects of didactic, modeling and experiential counselor training on reflection of feelings. <u>Dissertation Abstracts International</u>, 1972, <u>33</u>, 1276B.

Stoffer, D. L. An investigation of therapeutic success as a function of genuineness, nonpossessive warmth, empathic understanding, and dogmatism in the helping person. <u>Dissertation Abstracts International</u>, 1968, <u>29</u>, 799A.

Truax, C. B. The process of group psychotherapy: Relationships between hypothesized therapeutic conditions and intrapersonal exploration. <u>Psychological Monographs</u>, 1961, <u>75</u> (Whole No. 511), 1-35. (a)

Truax, C. B. Therapeutic conditions. <u>Psychiatric Institute Bulletin</u> (Wisconsin Psychiatric Institute), 1961, <u>1</u>, 10. (b)

Truax, C. B. Effective ingredients in psychotherapy: An approach to unraveling the patient-therapist interaction. <u>Journal of Counseling Psychology</u>, 1963, <u>10</u>, 256-263.

Truax, C. B. Influence of patient statements on judgments of therapist statements during psychotherapy. <u>Journal of Clinical Psychology</u>, 1966, <u>22</u>, 335-337. (a)

Truax, C. B. Reinforcement and nonreinforcement in Rogerian psychotherapy. <u>Journal of Abnormal Psychology</u>, 1966, <u>71</u>, 1-9. (b)

Truax, C. B. The meaning and reliability of accurate empathy ratings. <u>Psychological Bulletin</u>, 1972, <u>77</u>, 397-399.

Truax, C. B., & Carkhuff, R. R. For better or for worse . . . The process of psychotherapeutic personality change. In B. T. Wigdor (Ed.), <u>Recent advances in the study of behavior change</u>. Montreal: McGill University Press, 1963.

Truax, C. B., & Carkhuff, R. R. Significant developments in psychotherapy research. In L. E. Abt & B. F. Riess (Eds.), <u>Progress in clinical psychology</u>. New York: Grune and Stratton, 1964.

Truax, C. B., & Carkhuff, R. R. Experimental manipulations of therapeutic conditions. <u>Journal of Consulting Psychology</u>, 1965, <u>29</u>, 119-124.

Truax, C. B., & Carkhuff, R. R. <u>Toward effective counseling and psychotherapy: Training and practice</u>. Chicago: Aldine, 1967.

Truax, C. B., Carkhuff, R. R., & Douds, J. Toward an integration of the didactic and experiential approaches to training in counseling and psychotherapy. <u>Journal of Counseling Psychology</u>, 1964, <u>11</u>, 240-247.

Truax, C. B., Carkhuff, R. R., & Kodman, F. Relationship between therapist-offered conditions and patient change in group psychotherapy. Journal of Clinical Psychology, 1965, 21, 327-329.

Truax, C. B., & Lister, J. L. Effects of short-term training upon accurate empathy and non-possessive warmth. Counselor Education and Supervision, 1971, 10, 120-125.

Truax, C. B., & Mitchell, K. M. Research on certain therapist interpersonal skills in relation to process and outcome. In A. E. Bergin & S. L. Garfield (Eds.) Handbook of psychotherapy and behavior change. New York: John Wiley & Sons, 1971.

Truax, C. B., Wargo, D. G., & Silber, L. D. Effects of group psychotherapy with high accurate empathy and nonpossessive warmth upon female institutionalized delinquents. Journal of Abnormal Psychology, 1966, 71, 267-274.

Winett, R. A., & Winkler, R. C. Current behavior modification in the classroom: Be still, be quiet, be docile. Journal of Applied Behavior Analysis, 1972, 5, 449-504.

Yoakley, R. W. The effects of a traditional program compared with video and symbolic modeling, feedback and reinforcement as a means of teaching empathy to counselors for the deaf. Dissertation Abstracts International, 1972, 32, 4437A.

Footnotes

[1]Adapted by Paul Yelsma. For reprints write: Communications Department, Western Michigan University, Kalamazoo, Michigan.

[2]Their definition of empathy involves "emotional responsiveness to others' concerns" and getting "aroused by others' emotional experiences of both positive and negative qualities."

[3]Katz, R. "An Evaluation of the Volunteer Training Program at Gryphon Place." Unpublished Master's Thesis, Western Michigan University, 1976.