STOFFER, Dean Lewis, 1934-
AN INVESTIGATION OF THERAPEUTIC SUCCESS AS A FUNCTION OF GENUINENESS, NONPOSSESSIVE WARMTH, EMPATHIC UNDERSTANDING, AND DOGMATISM IN THE HELPING PERSON.

The Ohio State University, Ph.D., 1968
Education, guidance and counseling

University Microfilms, Inc., Ann Arbor, Michigan
AN INVESTIGATION OF THERAPEUTIC SUCCESS AS A FUNCTION OF GENUINENESS, NONPOSSESSIVE WARMTH, EMPATHIC UNDERSTANDING, AND DOGMATISM IN THE HELPING PERSON

DISSERTATION

Presented in Partial Fulfillment of the Requirements for the Degree Doctor of Philosophy in the Graduate School of The Ohio State University

By

Dean Lewis Stoffer, B.A., M.A.

The Ohio State University
1968

Approved by

Donald C. Smith
Adviser
Department of Education
ACKNOWLEDGMENTS

The writer wishes to express his gratitude to Dr. Donald C. Smith for his encouragement and direction during the period of this study and throughout the writer's doctoral program.

Appreciation is also expressed to Dr. Viola M. Cassidy and Dr. Herman J. Peters who read the manuscript and offered helpful suggestions.

Special thanks are also extended to Mrs. Loetta Hunt and Dr. Charles Huelsman whose faith in the writer provided needed moral support.

Finally, a sincere debt of gratitude is expressed to my wife for her faith and understanding.
VITA

October 17, 1934  Born - Mansfield, Ohio

1956 ........ B.A., North Central College, Naperville, Illinois

1956-1958 ... Mathematics Teacher, Lakeville Community High School, Otisville, Michigan

1959 ........ M.A., Ohio State University, Columbus, Ohio

1959-1961 ... School Counselor, Franklin Heights High School, Columbus, Ohio

1961 ........ Intern School Psychologist, Montgomery County Schools, Dayton, Ohio

1962-1965 ... School Psychologist, Delaware County Schools, Delaware, Ohio

1965-1966 ... H.E.W. Doctoral Fellow, The Ohio State University, Columbus, Ohio

1966 ........ Coordinator, Community Helper Program, Ohio State University, Columbus, Ohio

PUBLICATION

Stoffer, Dean L., A Panel Discussion of the Internship in School Psychology in Ohio. Published by the Lancaster City Schools and Division of Special Education, 1963.
<table>
<thead>
<tr>
<th>CONTENTS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>TABLES</td>
<td>vi</td>
</tr>
<tr>
<td>FIGURES</td>
<td>vii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Statement of the Problem</td>
<td></td>
</tr>
<tr>
<td>Assumptions</td>
<td></td>
</tr>
<tr>
<td>Limitations</td>
<td></td>
</tr>
<tr>
<td>Definition of Terms</td>
<td></td>
</tr>
<tr>
<td>Organization of the Study</td>
<td></td>
</tr>
<tr>
<td>II. REVIEW OF LITERATURE</td>
<td>11</td>
</tr>
<tr>
<td>Essential Ingredients of the Therapeutic</td>
<td></td>
</tr>
<tr>
<td>Relationship</td>
<td></td>
</tr>
<tr>
<td>Ratings of Therapeutic Conditions and</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td></td>
</tr>
<tr>
<td>Perceived Therapeutic Conditions and</td>
<td></td>
</tr>
<tr>
<td>Outcome</td>
<td></td>
</tr>
<tr>
<td>Relationship between Rated and Perceived</td>
<td></td>
</tr>
<tr>
<td>Conditions</td>
<td></td>
</tr>
<tr>
<td>The Question of Causation</td>
<td></td>
</tr>
<tr>
<td>The Open Mind and Provision of Therapeutic</td>
<td></td>
</tr>
<tr>
<td>Conditions</td>
<td></td>
</tr>
<tr>
<td>III. PROCEDURES</td>
<td>45</td>
</tr>
<tr>
<td>Selection and Description of Community</td>
<td></td>
</tr>
<tr>
<td>Helper Children</td>
<td></td>
</tr>
<tr>
<td>Selection and Description of Community</td>
<td></td>
</tr>
<tr>
<td>Helpers</td>
<td></td>
</tr>
<tr>
<td>The Instruments</td>
<td></td>
</tr>
<tr>
<td>Procedures for Collection of Data</td>
<td></td>
</tr>
<tr>
<td>Procedures for Analysis of Data</td>
<td>iv</td>
</tr>
</tbody>
</table>
### CONTENTS (Contd.)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>IV. STATISTICAL ANALYSIS AND INTERPRETATION OF DATA</td>
<td>66</td>
</tr>
<tr>
<td>Reliability of Ratings</td>
<td></td>
</tr>
<tr>
<td>Means, Range and Variability of Ratings</td>
<td></td>
</tr>
<tr>
<td>Stability of Rated Conditions</td>
<td></td>
</tr>
<tr>
<td>Characteristics of Relationship</td>
<td></td>
</tr>
<tr>
<td>Inventory Scores</td>
<td></td>
</tr>
<tr>
<td>Characteristics of Dogmatism Scores</td>
<td></td>
</tr>
<tr>
<td>Intercorrelations among Predictor Variables</td>
<td></td>
</tr>
<tr>
<td>Results of Multiple Correlations</td>
<td></td>
</tr>
<tr>
<td>Summary of Chapter IV</td>
<td></td>
</tr>
<tr>
<td>V. SUMMARY, CONCLUSIONS, AND IMPLICATIONS</td>
<td>88</td>
</tr>
<tr>
<td>Summary of Procedure</td>
<td></td>
</tr>
<tr>
<td>The Findings</td>
<td></td>
</tr>
<tr>
<td>Implications</td>
<td></td>
</tr>
<tr>
<td>Suggestions for Further Study</td>
<td></td>
</tr>
<tr>
<td>APPENDIXES</td>
<td>99</td>
</tr>
<tr>
<td>BIBLIOGRAPHY</td>
<td>128</td>
</tr>
<tr>
<td>Table</td>
<td>Page</td>
</tr>
<tr>
<td>-------</td>
<td>------</td>
</tr>
<tr>
<td>1. Reliabilities of Rating Scales for Accurate Empathy, Nonpossessive Warmth, and Genuineness</td>
<td>27</td>
</tr>
<tr>
<td>2. Means, Standard Deviations, and Reliability Coefficients for the Rokeach Dogmatism Scale (Form E)</td>
<td>56</td>
</tr>
<tr>
<td>3. Reliability Coefficients Among Three Judges in the Rating of Three Variables</td>
<td>67</td>
</tr>
<tr>
<td>4. Pearson Product-Moment Correlations between Rated Conditions—Early and Late</td>
<td>74</td>
</tr>
<tr>
<td>5. Pearson Product-Moment Intercorrelations Among Five Predictor Variables</td>
<td>78</td>
</tr>
<tr>
<td>6. Intercorrelation between Five Predictor Criterion and Each of Five Outcome Indexes</td>
<td>83</td>
</tr>
<tr>
<td>7. Multiple Correlation Using Five Selected Variables to Predict Therapeutic Outcome</td>
<td>86</td>
</tr>
<tr>
<td>Figure</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>------</td>
</tr>
<tr>
<td>1. Per Cent of Ratings Assigned to Each Stage of the Nonpossessive Warmth Scale in the Community Helper Program and the Wisconsin Study</td>
<td>69</td>
</tr>
<tr>
<td>2. Per Cent of Ratings Assigned to Each Stage of the Accurate Empathy Scale in the Community Helper Program and in the Wisconsin Study</td>
<td>71</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

The helping relationship (counseling or psychotherapy) is intended to produce constructive behavioral and personality change. Rapid growth over the years would suggest that clients as well as professionals are convinced of its value. Included among those who practice it are clinical psychologists, psychiatrists, social workers, school counselors, parole officers, clergymen, and physicians to name but a few.

However, considerable evidence suggests that counseling or psychotherapy is not superior to "no treatment." A report by Eysenck combined the reports of nineteen evaluations covering more than 7,000 cases of neurotic patients treated by eclectic or psychoanalytic approaches to counseling (1961). He concluded that:

With the single exception of the psychotherapeutic methods based on learning theory, results of published research with military and civilian neurotics, and with both adults and children, suggest that the therapeutic effects of psychotherapy are small or nonexistent and do not in any demonstrable way add to the non-specific effects of routine medical treatment or to such events as occur in the patient's everyday experience (p. 720).

Levitt (1957) examined eighteen evaluations of psychotherapy with children involving 3,399 cases and compared
their results with similar evaluations of untreated children. He concluded his report with, "The results of the present study fail to support the view that psychotherapy with neurotic children is effective" (p. 195). Thus, evidence involving large numbers of cases suggests that the average effects of therapeutic interventions are approximately equivalent to the random effects of normal living.

Despite the overwhelming evidence that psychotherapy, on the average, is no better than no treatment, specific studies involving specific therapists do demonstrate positive effects (Barron and Leary, 1955; Cartwright and Vogel, 1960; Truax, 1963; Mink and Issacson, 1959; and Teuber and Powers, 1953).

The reports suggest not only that some therapists are much more effective than others but also that some therapists are helpful while others are ineffective or harmful. The average effect is comparable to receiving no help.

After an examination of the literature, Truax and Carkhuff (1966) reported the following conclusions concerning the effectiveness of psychotherapy:

1. Certain relatively unspecified kinds are indeed effective.

2. Under certain unspecified conditions, therapy and control patients show equivalent outcomes, but those treated by psychotherapy show greater variability in outcome than those in control conditions.

3. What is called psychotherapy or counseling is a heterogeneous collection of ingredients or psychological conditions that produce varying degrees
of both positive and deteriorative personality change in patients (p. 21).

The conclusions based on the evaluations of counseling and psychotherapy suggest that psychotherapy research should be concerned with identifying the specific ingredients of the psychotherapeutic process which relate to client outcome. The question becomes, "What are the essential elements in the psychotherapeutic process which lead to constructive behavioral change in the client?"

Rogers (1957) described constructive personality change as

Change in the personality structure of the individual at both surface and deeper levels, in a direction which clinicians would agree means greater integration, less internal conflict, more energy utilizable for effective living; change in behavior away from behaviors generally regarded as immature and toward behaviors regarded as mature (p. 95).

The "necessary and sufficient" conditions for personality change include:

1. Two persons in a psychological relationship.

2. The first, called the client, is in a state of incongruence, being vulnerable or anxious. (Incongruence refers to a discrepancy between the actual experience of the organism and the self picture of the individual insofar as it represents the experience.)

3. The second person, the therapist, is congruent or integrated in the relationship.

4. The therapist experiences unconditional positive regard for the client.

5. The therapist experiences an empathic understanding of the client's internal frame of reference.
6. Communication of positive regard and empathic understanding (p. 96).

Rogers' statement has served as impetus to further thought and research and is the basis for the present research. While it is unlikely that the three conditions of congruence, unconditional positive regard, and empathic understanding will prove either necessary in all cases or sufficient in any case, the role of these three central conditions needs further exploration.

Kemp has suggested that highly dogmatic persons have relatively little communication between substructures of their personality (Kemp, 1961). They are described as less aware of the feelings and desires at the center of their being. They reject or distort ideas which do not comfortably fit into their belief system. Highly dogmatic persons could not, then, be expected to be highly genuine. In addition they could be expected to provide high levels of warmth only when the client appears to share the same belief system. Finally, unless one is able to understand and respect one's own belief system, it is unlikely that an individual would be able to understand and respect the belief system of another. Thus it can be predicted that the highly dogmatic person will be less likely to be highly genuine or to provide high levels of nonpossessive warmth or accurate empathy in the helping relationship and therefore be less effective.
Statement of the Problem

This study attempts to find support for the factors in the "helping person" that are hypothesized to bring about the therapeutic result.\(^1\) Based on the conclusions made from the evidence cited it has been predicted that the helping person who is genuine and who provides relatively high levels of nonpossessive warmth (unconditional positive regard) and accurate empathy (empathic understanding) will bring about more positive behavioral change in the person being helped than the helping person who is ungenuine or who provides relatively low levels of nonpossessive warmth and accurate empathy. In addition, an attempt is made to identify those helping persons who are most likely to bring about positive behavioral change.

The purpose of this study, then, is one of attempting to explore whether or not:

1. Positive behavioral change in children is related to genuineness, both rated and perceived, on the part of the helping person within the interpersonal relationship.

\(^1\)This research was a part of a larger project entitled "Utilization of Community Helpers for Meeting the Psychological and Educational Needs of Emotionally Handicapped Children"; a demonstration proposal supported by the Division of Handicapped Children and Youth, Department of Health, Education and Welfare, under provisions of public law 88-164.

The Community Helper Program attempted to determine the feasibility and usefulness of allowing untrained community volunteers to work individually with children who were both underachievers and behavior management problems. The period of investigation extended from September 1, 1966, to June, 1967.
2. Positive behavioral change in children is related to the receiving of relatively high amounts of nonpossessive warmth, both rated and perceived, from the helping person.

3. Positive behavioral change in children is related to the receiving of relatively high amounts of empathic understanding, both rated and perceived, from the helping person.

4. Persons relatively high in dogmatism are less likely to effect positive behavioral change in children than those lower in dogmatism.

Assumptions

For the purpose of this study it was assumed that

1. Positive behavioral change in elementary school age children is an observable phenomenon which can be measured.

2. Subjects would not change their normal interview behavior significantly when interviews were being audio tape recorded.

Limitations

For the purpose of this study the following limitations appear operative:

1. The results of this study should be applicable only when similar subjects are used and when procedures for measurement are nearly identical. Since much of the data is based upon objective ratings, use of identical scales will not guarantee similar measurements unless background and
training of raters are also similar to those of the present study.

2. Instruments used in this study have not been used extensively in similar situations. The rating scales and dogmatism scale have been utilized in other studies and have been shown to be useful. The relationship inventory, however, was created for this study. Its usefulness will be determined by its ability to predict therapeutic outcome.

**Definition of Terms**

Truax and Carkhuff (1966) have explained the meaning of accurate empathy, nonpossessive warmth, and genuineness in the following manner:

Genuineness. This scale is an attempt to define five degrees of therapist genuineness, beginning at a very low level where the therapist presents a facade or defends and denies feelings; and continuing to a high level of self-congruence where the therapist is freely and deeply himself. A high level of self-congruence does not mean that the therapist must overtly express his feelings but only that he does not deny them. Thus, the therapist may be actively reflecting, interpreting, analyzing, or in other ways functioning as a therapist; but this functioning must be self-congruent, so that he is being himself in the moment rather than presenting a professional facade. Thus the therapist's response must be sincere rather than phony; it must express his real feelings or being rather than defensiveness.

"Being himself" simply means that at the moment the therapist is whatever his response denotes. It does not mean that the therapist must disclose his total self, but only that whatever he does show is a real aspect of himself, not a response growing out of defensiveness or a merely "professional" response that has been learned and repeated (p. 68).
Nonpossessive Warmth. The dimension of nonpossessive warmth or unconditional positive regard, ranges from a high level where the therapist warmly accepts the patient's experience as part of that person, without imposing conditions; to a low level where the therapist evaluates a patient or his feelings, expresses dislike or disapproval or expresses warmth in a selective and evaluative way.

Thus, a warm positive feeling toward the client may still rate quite low in this scale if it is given conditionally. Nonpossessive warmth for the client means accepting him as a person with human potenti-
alities. It involves a nonpossessive caring for him as a separate person and, thus, a willingness to share equally his joys, and aspirations or his depressions and failures. It involves valuing the patient as a person, separate from any evaluation of his behavior or his thoughts but still rate high on warmth if it is quite clear that his valuing of the individual as a person is unconditional and uncontaminated. At its highest level this unconditional warmth involves a nonpossessive caring for the patient as a separate person who is allowed to have his own feelings and experiences; a prizing of the patient for himself regardless of his behavior.

It is not necessary--indeed, it would seem undesirable--for the therapist to be nonselective in reinforcing, or to sanction or approve thoughts and behaviors that are disapproved by society. Nonpossessive warmth is present when the therapist appreciates such feelings or behaviors and their meaning to the client, but shows a nonpossessive caring for the person and not for his behavior. The therapist's response to the patient's thoughts or behaviors is a search for their meaning or value within the patient rather than disapproval or approval (p. 58).

Accurate Empathy. Accurate empathy involves more than just the ability of the therapist to sense the client's or patient's "private world" as if it were his own. It also involves more than just his ability to know what the patient means. 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.
It is not necessary—indeed it would seem undesirable—for the therapist to share the client's feelings in any sense that would require him to feel the same emotions. It is instead an appreciation and a sensitive awareness of those feelings. At deeper levels of empathy, it also involves enough understanding of patterns of human feelings and experience to sense feelings that the client only partially reveals. With such experience and knowledge, the therapist can communicate what the client clearly knows as well as meanings in the client's experience of which he is scarcely aware.

At a low level of accurate empathy the therapist may go off on a tangent of his own or may misinterpret what the client is feeling. At a very low level he may be so preoccupied and interested in his own intellectual interpretations that he is scarcely aware of the client's "being." The therapist at this low level of accurate empathy may even be uninterested in the client, or may be concentrating on the intellectual content of what the client says rather than what he "is" at the moment, and so may ignore or misunderstand the client's current feelings and experiences. At this low level of empathy the therapist is doing something other than "listening," "understanding," or "being sensitive"; he may be evaluating the client, giving advice, sermonizing, or simply reflecting upon his own feelings or experiences. Indeed, he may be accurately describing psychodynamics to the patient—but in the wrong language for the client, or at the wrong time, when these dynamics are far removed from the client's current feelings, so that the interaction takes on the flavor of "teacher-pupil" (p. 46).

Rokeach (1960) has defined the open belief system.

Every person, then must be able to evaluate adequately both the relevant and irrelevant information he receives from every situation. This leads us to suggest a basic characteristic that defines the extent to which a person's system is open or closed; namely, the extent to which a person can receive, evaluate, and act on relevant information received from the outside on its own intrinsic merits, unencumbered by irrelevant factors in the situation arising from within the person or from the outside. . . . The more open one's belief system, the more should evaluating and acting on information proceed independently on its own merits, in accord with the
structural requirements of the situation. Also the more open the belief system the more should the person be governed in his actions by internal self-actualizing forces and the less by irrational inner forces (p. 57).

**Organization of the Study**

The present chapter includes an introduction to the subject of investigation, a review of certain relevant literature, a statement of the problem to be investigated, assumptions, limitations pertinent to its validity, and definitions of terms.

The second chapter contains a more comprehensive review of related literature including research findings. The third chapter describes the framework from which the study has developed, the population of the study, the instruments used in measuring the independent variables, and procedures used for collection and analysis of data.

The fourth chapter contains the statistical analysis and interpretation of the data. The final chapter contains a summary of the data presented in this study, conclusions based upon the data, and implications of the findings. Recommendations are made for further investigation.
CHAPTER II

REVIEW OF LITERATURE

The hypotheses to be examined in this study deal with three interpersonal conditions and their relation to positive behavioral change. In addition, it has been predicted that the dogmatic person is less likely to provide these three conditions and therefore will be less successful in the therapeutic relationship than the open minded person. The review of literature concerning the above areas will be presented in six sections. These topics include (1) essential ingredients of the therapeutic relationship; (2) ratings of therapeutic conditions and outcomes; (3) perceived therapeutic conditions and outcome; (4) relationship between ratings and perceptions of therapeutic conditions; (5) the question of causation; and (6) helper dogmatism and provision of therapeutic conditions.

Essential Ingredients of the Therapeutic Relationship

Studies by Fiedler have given considerable impetus to the quest for conditions common to all therapeutic relationships. Using a Q-technique, he had eight therapists (four psychoanalytic, two nondirective, two eclectic) describe the
relationship which they considered ideal (Fiedler, 1950a). All therapists correlated positively with each other (.48 to .78). However, he found that responses of experienced therapists from different schools correlated more highly with each other than with less experienced therapists from the same school ($p < .05$). There were no individual items which emerged as major differences between schools. Those items most likely to be considered as therapeutic included:

- an empathic relationship
- therapist and patient relate well
- therapist sticks closely to the patient's problems
- patient feels free to say what he likes
- atmosphere of mutual trust and confidence exists
- rapport is excellent
- patient assumes an active role
- therapist leaves patient free to make his own choices
- therapist accepts all feelings which the patient expresses as completely normal and understandable
- a tolerant atmosphere exists
- an understanding therapist
- patient feels most of the time that he is really understood
- therapist is really able to understand patient
- the therapist really tries to understand patient's feelings (p. 241)

A second study supported the results of the first. Fiedler concluded that "theoretical allegiance to one system of
psychotherapy does not change the therapist's goal with respect to the relationship which he strives to create with his patient. . . . it is expertness which determines the type of relationship which is set as a goal by therapist" (p. 244).

To further explore his hypotheses, Fiedler used ten recorded therapy interviews (four analytic, four nondirective, and two Adlerian) including at least one expert and one novice in each group (Fiedler, 1950). One untrained and three trained judges were asked to describe the relationship by use of a 75 item Q-technique. All three nationally known experts correlated more highly with the ideal than any of the nonexperts. Nonexperts and experts of the same school were clearly separated.

Fiedler was specifically interested in three dimensions: communication, emotional distance, and status. He concluded that ability to understand the patient was the most important of the criteria of expertness. Involvement and status were not significantly related to the criteria.

In a final study, Fiedler (1951) factor analyzed the statements of his four judges. He found no factors which characterized therapists of any of the three schools. The most characteristic aspects of the relationship which differentiated experts from nonexperts were their greater ability to understand the feelings of the patient, greater security in the therapeutic situation and capacity to show interest and warmth without becoming overly involved with the patient.
Strupp, Wallach, and Wogan (1964) reported on a questionnaire survey of therapists and patients which indicated consensus about essential aspects and outcome of psychotherapy. They summarize their findings as follows:

The emergence of a large general "warmth" factor deserves emphasis. There was additional evidence to suggest that overshadowing this attitudinal-emotional factor is the patient's conviction that he has the therapist's respect. This faith in the integrity of the therapist as a person may be called the capstone of a successful therapeutic relationship subsuming other characteristics. Technical skill on the part of the therapist may go a long way to capitalize on such a relationship, although the present data does not specifically inform us how such a relationship comes into being, is deepened, and tuned into maximum therapeutic advantage. However, there is little doubt that a relationship having these qualities represents the most basic ingredient of beneficial therapeutic influence irrespective of the formal aspects of the setting (p. 35).

Covering a decade, Whitehorn and Betz (1954) analyzed the differences between physicians and their different styles of transactions with schizophrenic patients' progress and outcome. Seven successful and seven unsuccessful physicians and 100 patients made up the populations for the study. Information on the therapists and patients was taken from individual case records.

The investigators made the following conclusions:

We interpret these empirical findings to mean that in the psychotherapy of schizophrenic patients success is to a large extent determined by the differences found among physicians in the extent to which they are able to approach their patient's problems in a personal way, gain a trusted, confidential relationship and participate in an active, personal way in the patient's reorientation to personal relationships. Techniques of passive permissiveness, or efforts to develop insights by
interpretation appear to have much less therapeutic value (Whitehorn and Betz, 1954, p. 331).

Studies by Feidler (1950, 1950a, 1951); Strupp, Wallach, and Wogan (1964); and Whitehorn and Betz (1954) have presented evidence which suggests that there are common elements in the therapeutic relationship which cut across schools and theories of psychotherapy. Warmth, understanding, trust, and respect emerge as basic to these common elements.

Ratings of Therapeutic Conditions and Outcome

Publication of the "Necessary and Sufficient Conditions of Therapeutic Personality Change" enabled researchers to test the hypotheses empirically (Rogers, 1957). Halkides (1958), a student of Rogers, provided supporting evidence through the use of rating scales. The purpose of the study was to relate positive behavioral change to empathic understanding in the therapist, unconditional positive regard in the therapist, congruence in the therapist, and high relationship between affective intensity of corresponding therapist responses.

An index of success was determined by counselor rating of the outcome of therapy, TAT ratings of adjustment, and self-ideal correlations of a Q-sort. The data was taken from 20 cases picked randomly from a research group at the University of Chicago. Using the multiple criteria, ten cases were considered more successful and ten were considered less successful. Two interviews were picked from each case; one
early interview and one near the end of therapy. From each of these 40 interviews, nine random units of client statement and the corresponding client response were transferred to a tape which made up the basic data of the investigation. Three judges listened to the recorded material and with the aid of the typescripts and the scales made ratings of the operationally defined variables.

The results were checked for reliability of judges' ratings and found to be high (empathic understanding, .99; unconditional positive regard, .98; congruence, .91). Affective intensity reliabilities were not high enough to merit combination. The findings gave strong support for the importance of empathic understanding, unconditional positive regard, and congruence in differentiating more and less successful cases. All were significant beyond the .001 level. Further analysis indicated a high relationship between variables (empathic understanding—unconditional positive regard, r=.89; empathic understanding—congruence, r=.72; unconditional positive regard—congruence, r=.78).

A study by Hart (1960) entitled "A Replication of the Halkides Study" failed to confirm these results.

The Truax studies

Following the example set by Halkides, Truax and Carkhuff (1966) have conducted a series of studies in which they developed and used the rating scales selected for the present study. Much of the evidence has come from the study
of psychotherapy with 16 hospitalized schizophrenic patients begun in 1958 at the University of Wisconsin.

In 1963, Truax summarized the findings to that time. In an initial study involving the accurate empathy scale, four patients who showed clear improvement and four patients who showed deterioration on a battery of psychological tests after six months of therapy were selected. Three-hundred eighty-four samples of tape-recorded psychotherapy were randomly selected and coded so that raters would not know whether a sample came from an improved or a deteriorated case, or from an early or late interview. The psychotherapy involving test-improved patients rated consistently higher on accurate empathy than did the psychotherapy with test-deteriorated cases ($p < .01$).

The next study using the accurate empathy scale involved 14 hospitalized schizophrenic cases and 14 counseling cases from the University of Chicago and Stanford University (Truax, 1963). Analysis of 112 samples of recorded psychotherapy indicated that accurate empathy ratings were significantly higher for the more successful cases than for the less successful ($p < .01$). The positive relationship between accurate empathy and outcome held for both hospitalized schizophrenics and counseling cases.

In still another study dealing with the effects of accurate empathy on schizophrenics, the time span was increased to three and one-half years (Truax, 1963). One
four-minute tape recorded sample was taken from every fifth interview for each of 14 cases. Over the extended period of time, therapists showed no tendency to systematically change in the level of accurate empathy offered to the patient. Therapists of the more improved patients were judged to have offered significantly higher levels of accurate empathy throughout the program than did therapists of the unimproved patients.

Using the same procedures, the scales for nonpossessive warmth and therapist genuineness were applied to the data involving the 14 schizophrenic patients (Truax, 1963). It was found that in improved cases, therapists were consistently rated higher in nonpossessive warmth and therapist genuineness than in nonimproved or failure cases (p. < .05).

The investigators then compared therapy cases rated high on all three conditions with patients receiving no therapy and therapy cases rated low on all three ingredients (Truax, 1963). Using the ratings already available, they computed the mean value for total therapeutic conditions. Six patients were judged to have received relatively high levels of therapeutic conditions while eight patients were judged to have received relatively low levels. Pre and post test batteries were analyzed by two clinical psychologists unfamiliar with the program. The Rorschach and MMPI were used in assessing psychological functioning. Patients receiving high levels of therapeutic conditions showed an
over-all gain whereas patients who received relatively low levels of therapeutic conditions showed a loss in psychological functioning. Control patients evidenced moderate gains. The investigators concluded that the data suggests that the high therapeutic conditions facilitate positive behavioral change and that low conditions promote negative behavioral change.

Additional analysis of the same data dealt with the relation between conditions offered in therapy and change in anxiety experienced by the patient (Truax, 1963). Data from the Anxiety Reaction Scale indicated a tendency for patients receiving high levels of therapeutic conditions to show a drop in anxiety level while patients receiving low levels of conditions evidenced an increase in anxiety level. Control cases showed no significant change.

Application of a Q-sort for self data found that the self-concept of patients receiving high levels of conditions showed a slight tendency toward adjustment from early to later interviews, while patients receiving low levels of therapy showed marked change toward a less well-adjusted self-concept (Truax, 1963).

Finally, the investigators used a Constructive Personality Change Index derived from items from early and late MMPI tests (Truax, 1963). On this index, the high conditions therapy patients showed large positive changes in personality functioning; the low conditions therapy patients showed
moderate negative gains; and the control group patients showed moderate positive changes.

In recent studies, Truax and associates have attempted to cross-validate earlier evidence (Truax, Wargo, Frank, Imber, Battle, Hoehn-Saric, Nash, and Stone, 1966). The population was made up of 40 out-patients treated by resident psychiatrists at the Phipps Psychiatric Clinic at John Hopkins Hospital. Each four psychiatrists worked with ten patients. Therapists met with patients at least once a week for one-hour sessions and all cases were terminated at four months. The three scales developed by Truax were used as the basic measures of conditions. Four judges rated each scale after training. Reliability as measured by intraclass correlation for combined judges in the mean ratings per case were: therapist accurate empathy, \( R = 0.63 \); therapist nonpos­sive warmth, \( R = 0.59 \); and therapist genuineness, \( R = 0.60 \).

Measures of total improvement included patient global improvement scales filled out by the therapist, patient global improvement scales filled out by patient, social ineffectiveness ratings filled out after therapy by a research interviewer, and target symptom improvement scales filled out by the patient after therapy.

Degree of improvement, in general, was not related to the initial level of adjustment. As predicted, patients who received the highest levels of empathy, warmth, and genuineness combined tended to show significantly greater improvement
in the two global measures. Differences on the three more specific measures of patient improvement did not reach significance.

The investigators concluded that the over-all findings tended to support the hypotheses predicting that patients in therapy with psychiatrists offering high levels of therapeutic conditions show greater improvement than patients in psychotherapy with therapists providing relatively lower conditions. They conclude that psychotherapy can be for better or worse, depending upon the levels of conditions provided during treatment. In addition, differences in outcome even when level of conditions appear to be identical suggest that other unknown factors also contribute toward outcome.

The Wisconsin study

In an elaborate series of studies conducted at the University of Wisconsin and reported by Rogers, Gendlin, Kiesler, and Truax (1967), the investigators struggled with many of the problems and goals of the present study. The purpose of the Wisconsin study was to investigate the relationship between the therapeutic triad—nonpossessive warmth, accurate empathy and genuineness—as well as process, process movement, and outcome in psychotherapy with schizophrenic adults. Each of eight therapists who had volunteered for the study was assigned one more chronic schizophrenic, one more
acute schizophrenic, and one normal individual. Apparently, these are the same 16 schizophrenic patients included in the Truax studies. In addition, 24 controls matched by sex, age, and socio-educational level were identified. The research data consisted of test batteries and research instruments given at intervals throughout the program, recorded therapy interviews held with experimental subjects, and recorded sampling interviews held at three-month intervals throughout the program with all subjects.

Data from pilot studies suggested that one to three samples per interview, two to five minutes in length, would give an adequate basis for ratings. Bocchini, Farwell, and Hart (1960) found that three to five minute segments yielded more reliable ratings than shorter segments. In a study comparing patient process ratings in segments of two minutes, four minutes, eight minutes, and 16 minutes, Kiesler, Mathieu, and Klein (1964) found no difference in range or reliability of ratings between the longer segments and the shorter. The majority of the segments used in the Wisconsin study were of four minutes duration and extracted at random from the latter half of the interview tape according to a standardized procedure.

Raters were undergraduates drawn from the student body of the University of Wisconsin. The investigators felt that clinically naive raters would be freer to adopt the set necessary for rating conditions from the recipient's viewpoint than would the expert clinician who would be especially attuned
to the intent and motivation of the therapist. Studies by Arnhoff (1954) and Cronbach (1960) suggest that naive assessors are often more reliable in their judgments and more likely to confine their judgments to the dimension at issue.

Each rater was trained in the use of only one scale and was not allowed to become acquainted with other scale dimensions. This was done to eliminate any "halo effects" between variables. The investigators found that the more structured approach to rater training produced more reliable ratings than did informally supervised practice.

Tape recorded segments were presented to all judges in a standard random order. Raters had no information concerning the nature of the case rated, the type of interview, the location of a given interview in the over-all course of therapy, or the outcome of the case.

Rated material consisted of selected segments from interviews two through 15 and every fifth interview thereafter for each of 14 experimental patients. Three groups of judges made independent ratings for each of the three conditions scales on a total of 337 segments. The score for each segment consisted of the average of the ratings by the three judges for that segment. The mean of all interviews rated were used for further analyses.

Intraclass reliabilities were computed according to Ebel's formula (Guilford, 1954), yielding an estimate of the reliability of the means of the judges' ratings and an
estimate of the average intercorrelation of all judge com-
binations. Good intraclass reliabilities were obtained for
the genuineness scale (r=.53). Reliability of accurate empa-
thy ratings were somewhat less but considered adequate
(r=.36). Nonpossessive warmth scale ratings were considered
to be unreliable (r=.08) and excluded from further analyses.

The investigators found little relative consistency
in the level of accurate empathy offered by the therapist
over the first few interviews. It was not until interviews
seven and eight that the accurate empathy in one hour could
be predicted from that of the preceding hour. Interviews
four, five, and nine were particularly atypical. Inter-
relationship among the last five interviews was particularly
high, suggesting that accurate empathy reached relative
stability as the relationship progressed. Genuineness ratings
were less stable than accurate empathy ratings over inter-
views two to 15. Relationship between successive interviews
tended to be uniformly positive but low.

The original outcome battery included the Minnesota
Multiphasic Inventory (MMPI), the Rorschach, the Thematic
Apperception Test, the Wechsler Adult Intelligence Scale,
the California F Test, the Wittenborn Psychiatric Rating
Scales, the Stroop Interference Test, the Butler-Haigh
Q-sort, the Truax Anxiety Scale, and the Therapist Rating
Scale. Whenever possible, all self-report measures were
administered at six-month intervals, with Wittenborn ratings
made every three months. For the purpose of analysis, only those test batteries yielding fairly complete coverage of experimental and control patients administered before and after the program were used. With these restrictions, only the MMPI, the Q-sort, and the Wittenborn instruments were used as indexes of personality and behavioral change in addition to therapist ratings and hospitalization indexes.

Support for the hypothesis that the experimental group would improve more than the control group was quite limited. Only hospitalization rates one year after therapy termination significantly favored experimental patients. In all other respects, the outcome-change measures for experimental and control groups were virtually identical.

Rated accurate empathy was used to differentiate high and low conditions for further analysis. When the experimental group was dichotomized according to the level of conditions received in the therapy relationship, there was a statistically significant positive relationship between level of conditions and measures of patient outcome. Patients who received high levels of accurate empathy showed positive change for the MMPI schizophrenia subscale while patients who received low levels of accurate empathy showed a slightly negative trend. The finding was supported by other consistent but statistically insignificant trends. In general, the control cases had a moderate level of improvement; the patients who received higher levels of therapeutic
conditions showed much improvement; and the patients who received low levels of conditions did not improve. The investigators concluded that while the results were not sufficiently reliable to lend unequivocal support to their hypothesis, they were sufficiently consistent to warrant further study.

**Discussion**

Reliability between raters, a necessary component in measurement of the therapeutic variables shows some variation. Halkides (1957), using relatively sophisticated judges, reported high levels of inter-judge reliability for all three variables. Table 1 presents reliability coefficients for several studies, using judges with varying degrees of training and experience. From this table, it would appear that judges were able to agree fairly well on all three variables for the Truax studies. The studies by Bergin and Solomon (1963) and Melloh (1964) were concerned with accurate empathy only. The Wisconsin Study (Rogers et al., 1967), using undergraduate college students as judges, reported adequate reliabilities between judges for accurate empathy and genuineness. However, inter-judge reliability was so low for nonpossessive warmth that it had to be excluded from further analysis.

Halkides, using the same three judges for each variable, reported a high relationship between the three rated
### TABLE 1

**RELIABILITIES OF RATING SCALES FOR ACCURATE EMPATHY, NONPOSSESSIVE WARMTH, AND GENUINENESS**

<table>
<thead>
<tr>
<th>Study</th>
<th>N Samples</th>
<th>N Patients</th>
<th>N Therapists</th>
<th>Accurate Empathy</th>
<th>Nonpossessive Warmth</th>
<th>Genuineness</th>
</tr>
</thead>
<tbody>
<tr>
<td>Truax (1961)</td>
<td>384</td>
<td>8</td>
<td>7</td>
<td>.87</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truax &amp; Carkhuff (1963)</td>
<td>297</td>
<td>14</td>
<td>10</td>
<td>.89</td>
<td>.50a</td>
<td>.40a</td>
</tr>
<tr>
<td>Truax &amp; Carkhuff (1963)</td>
<td>112</td>
<td>28</td>
<td>24</td>
<td>.69a</td>
<td>.55a</td>
<td></td>
</tr>
<tr>
<td>Truax (1962)</td>
<td>448</td>
<td>14</td>
<td>10</td>
<td>.69</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bergin &amp; Solomon (1963)</td>
<td>28</td>
<td>28</td>
<td>18</td>
<td>.79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melloh (1964)</td>
<td>56</td>
<td>28</td>
<td>28</td>
<td>.62</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Truax, Wargo, Frank, Imber, Battle,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hoehn-Saric, Nash, &amp; Stone (1966a)</td>
<td>182</td>
<td>40</td>
<td>4</td>
<td>.63</td>
<td>.59</td>
<td>.60</td>
</tr>
<tr>
<td>Wargo (1962)</td>
<td>297</td>
<td>14</td>
<td>10</td>
<td>.89</td>
<td>.50</td>
<td></td>
</tr>
<tr>
<td>Truax &amp; Carhuff (1963)</td>
<td>64</td>
<td>8</td>
<td>8</td>
<td>.57a</td>
<td>.62a</td>
<td>.45a</td>
</tr>
<tr>
<td>Truax (1962a)</td>
<td>104</td>
<td>26</td>
<td>1</td>
<td>.69a</td>
<td>.55a</td>
<td>.40a</td>
</tr>
<tr>
<td>Truax, Wargo, Frank, Imber,</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Battle, Hoehn-Saric, Nash, &amp; Stone (1966b)</td>
<td>80</td>
<td>40</td>
<td>2</td>
<td>.75a</td>
<td>.57a</td>
<td>.55a</td>
</tr>
<tr>
<td>(Therapy Interviews)</td>
<td>182</td>
<td>40</td>
<td>4</td>
<td>.63</td>
<td>.59</td>
<td>.60</td>
</tr>
<tr>
<td>Truax (1966b)</td>
<td>50</td>
<td>5</td>
<td>5</td>
<td>.66a</td>
<td>.84a</td>
<td></td>
</tr>
<tr>
<td>(Edited)</td>
<td>50</td>
<td>5</td>
<td>5</td>
<td>.76a</td>
<td>.81a</td>
<td></td>
</tr>
<tr>
<td>Truax &amp; Carkhuff (1965a)</td>
<td>45</td>
<td>3</td>
<td>1</td>
<td>.78</td>
<td>.70</td>
<td>.83</td>
</tr>
<tr>
<td>Carhuff &amp; Truax (1965)</td>
<td>151</td>
<td>70</td>
<td>28</td>
<td>.43a</td>
<td>.48a</td>
<td>.62a</td>
</tr>
<tr>
<td>Truax &amp; Silber (1966)</td>
<td>144</td>
<td>48</td>
<td>16</td>
<td>.54</td>
<td>.52</td>
<td>.46</td>
</tr>
</tbody>
</table>

a. Average Pearson correlations. All others are Ebel intraclass reliabilities for the pooled data used in analysis of findings.
variables. She suggests that this indicates that they are related to an underlying basic attitude of the therapist that is common to all three variables. From the earlier studies, Truax and Carkhuff (1966) concluded that the achievement of a high degree of accurate empathy is dependent first on obtaining at least a minimally high level of nonpossessive warmth for the patient. In addition, they concluded that neither of these two conditions could function properly without the therapist being himself integrated and genuine within the therapeutic relationship. In a related study, Gendlin and Geist (1962) concluded that extremely low genuineness in therapy may invalidate the effects of higher levels in other conditions. However, evidence from the later Truax studies conflicted with that of earlier studies. In the later studies, the investigators found a high positive correlation between levels of accurate empathy and genuineness but nonpossessive warmth did not appear to be related to either of the other two. Rogers et al. (1967) were hampered in their attempt to explore the interrelationship of the conditions factors because of the questionable validity of the genuineness and nonpossessive warmth ratings. They found no consistent pattern of association among the three conditions factors.

Halkides (1957) found a strong positive correlation between each of the three therapeutic variables and case outcome. However, an attempt by Hart (1960) to replicate
the study was unable to confirm the results. In all of the studies reported by Truax and Carkhuff (1966), positive relationships were found between case outcome and the three therapeutic variables. Rogers et al. (1967) were less successful but did find a significant positive relationship between the level of accurate empathy provided by the therapist and outcome as measured by the MMPI schizophrenia subscale. Bergin and Solomon (1963) also found support for the hypothesis that accurate empathy in therapy was significantly related to therapeutic outcome.

Perceived Therapeutic Conditions and Outcome

Barrett-Lennard (1962) applied a different approach to the "necessary and sufficient" conditions. Since the client's experience of his therapist's response is the primary focus of therapeutic influence, he concluded that measurement should be made from the client's frame of reference. He predicted that empathic understanding, level of regard, unconditionality of regard, congruence, and willingness to be known are the five distinguishable aspects of the therapist's attitudes and responses which are influential in the process of therapeutic change.

For the purpose of evaluation, the Relationship Inventory was developed to yield measures of each of the five variables. Parallel forms were used for client and therapist. The Relationship Inventory consists of 92 items using three
degrees of agreement (+1,+2,+3) and three levels of disagreement (-1,-2,-3).

The subjects for the Barrett-Lennard study consisted of 42 clients who were participating in a large scale study of behavioral change associated with therapy. Twenty-one therapists worked with from one to four clients each. Length of therapy ranged from seven to 96 interviews, with a mean of 33 interviews.

Clients and therapists completed the Relationship Inventory after the first five interviews, after 15 interviews, after 25 interviews, and at the termination of therapy. Indexes of behavioral change at pre-therapy and post-therapy testing points included the therapist's ratings of client adjustment; self-descriptive data (Q-sort adjustment by Dymond, 1954); MMPI data; and the Taylor Manifest Anxiety Scale.

Internal reliability coefficients of each of the five individual scales were estimated using the Spearmen-Brown formula. All correlation coefficients were .82 or higher indicating considerable consistency. Therapists tended to be more consistent than clients and they also tended to describe the relationship more positively than did the clients. There was no evidence of change in the average quality of the client-therapist relationship at different points in therapy.
Pearson product-moment intercorrelations of the measures of relationship from client and from therapist perceptions after five interviews varied considerably. Correlation between accurate empathy and genuineness stood out markedly as the strongest relationship. The most independent variable in terms of client perceptions was unconditionality of regard. Correlations between the same dimensions measured from the two different vantage points tended to be low and insignificant.

Sixteen cases were considered more successful and 19 cases were considered less successful. As perceived by the clients, level of regard (p < .05), accurate empathy (p < .005), genuineness (p < .005), and unconditionality of regard (p < .01) were related to behavioral change as predicted. Willingness to be known was not significantly related. Therapist perceptions were also related in the predicted direction but less strongly. Therapists thought to be more competent scored higher on the four variables related to outcome than did therapists identified as less competent.

When the relationship evidence from client and therapist was combined, it provided a stronger, more discriminating prediction of assessed change than evidence from one or the other source. Finally, there was a slightly better client-therapist agreement with the expert therapists, especially in the area of accurate empathy. The investigator
concluded that the expert communicates accurate empathy more ambiguously and can identify it more accurately.

In addition to the rating scales, Rogers et al. (1967) used the Barrett-Lennard Relationship Inventory to assess the level of therapist conditions as perceived by patients and therapists. The first assessment reported was made shortly after the initiation of the therapy relationship. The second came near the final interview for each case. A sample of eight and 13 respectively for the initial and terminal patient Relationship Inventories and a sample of 11 and 14 for initial and terminal therapist assessments were obtained.

Results of rank-order correlations suggest that both patient and therapist perceptions of the therapeutic relationship were highly stable over the course of therapy. Correlations among patient Relationship Inventory subscales revealed positive but usually insignificant relationships. Each separate subscale correlated positively with the total score with only the accurate empathy subscale failing to reach the .05 level of significance. Intercorrelations among subtest scales for therapist Relationship Inventories were not reported.

Rank order correlations between subscales of patient and therapist Relationship Inventories at initial and terminal points in therapy were insignificant and tended to be negative. Therapists who rated themselves as relatively
high on the factors of genuineness and level of regard were seen by patients as relatively low in empathy and in conditions generally. In general, the therapist perceived himself as offering a relatively higher level of conditions than that perceived by or communicated to the patient. These results are similar to those reported by Barrett-Lennard although Barrett-Lennard did find a slight positive trend between patient and therapist perceptions.

Using the same data, Spotts (1962) found that in successful cases, there was a significant positive correlation between patient and therapist evaluations of positive regard, while in less successful cases, therapist and patients are more divergent, correlating negatively with one another.

Patients who perceived their therapists as highly genuine showed significant favorable hospitalization rates. This evidence was supported by other consistent but statistically insignificant trends. As was the case with interviews rated high in accurate empathy, patient perceived genuineness was related to the greatest gains. The control patients made moderate gains while the low conditions patients showed lack of improvement.

Discussion

Barrett-Lennard (1962) reported considerable consistency for each of the five individual scales of the Relationship Inventory (correlation of .82 or higher). Therapists were more consistent and more optimistic than clients. Both
clients and therapists were also consistent in their evaluation of the relationship over three inventory administrations. These findings were consistent with those reported in the Wisconsin Study (Rogers et al., 1967).

Barrett-Lennard (1962) found considerable variation in the correlations between subscales of the Relationship Inventory. In general, the correlations between therapist and client subscales tended to be positive but low and insignificant. The strongest single relationship was between accurate empathy and genuineness. In the Wisconsin Study (Rogers et al., 1967), as in the study reported by Barrett-Lennard, correlations between subscales of patient and therapist Relationship Inventories were insignificant. In contrast to the Barrett-Lennard study, however, they tended to be negative. Using the data from the Wisconsin study, Spotts (1962) found a significant positive correlation between patient and therapist evaluations in the area of positive regard for the more successful cases.

As perceived by clients, level of regard, accurate empathy, genuineness, and unconditionality of regard were found to be significantly related to positive behavioral change in the Barrett-Lennard study (1962). Therapist perceptions were also in the predicted direction but were nonsignificant. Combination of therapist and client perceptions provided a stronger, more discriminating prediction of positive change than either one alone. In the Wisconsin study
(Rogers et al., 1967), patient perceptions and especially patient perceived genuineness tended to be positively related to outcome. Therapist perceptions were not found to be related to outcome.

**Relationship between Rated and Perceived Conditions**

Only two studies have been reported which used both the rating scales and the Relationship Inventory. Truax (1966) compared three different variations of the rating technique with patient Relationship Inventories. Subjects included 40 hospitalized schizophrenics and 40 institutionalized male juvenile delinquents who participated in group psychotherapy for 24 sessions. The Relationship Inventory intercorrelated with the other measures less than the other three intercorrelated with each other. The investigators concluded that the Relationship Inventory was measuring different aspects of the therapy process than were the rating scales. The Relationship Inventory was also less predictive of outcome than the three rated conditions measures. The investigators hypothesized that since inaccurate interpersonal perceptions are common to emotional disturbance in most patient populations, patient inventories may be the poorest measures of therapist-offered conditions, and the relation between actual conditions and outcome may be independent of the patient's reported perceptions.
In the Wisconsin study (Rogers et al., 1967), the Relationship Inventory was administered shortly after the initiation of therapy and near the final interview for each case. Initially, the rated level of accurate empathy was highly correlated with all but the unconditionality subscale and most highly associated with empathy ($p < .01$) and total conditions on patient's Relationship Inventories ($p < .01$). Later in therapy, it was the patient's perception of therapist congruence that was most strongly associated with rated accurate empathy ($p < .01$). The investigators concluded that "at any point in therapy judges' ratings of therapist accurate empathy agreed significantly with the patient's assessments of therapist empathy and congruence as well as the over-all level of conditions offered in the therapeutic relationships" (p. 174).

Examination of the ratings of nonpossessive warmth where raters were able to agree also suggested a high association between these and most of the Relationship Inventory subscales and with the total score on the Relationship Inventory. Judges' ratings of therapist genuineness did not appear to be associated with patient perceptions. The investigators concluded that with the exception of the genuineness scale, raters and patients were in agreement regarding the gross level of therapist conditions. In contrast, therapists' Relationship Inventory evaluations and judges' ratings of the relationship were either not associated or negatively associated.
Discussion

Evidence concerning the relationship between rated and perceived conditions is small and inconclusive. Truax (1966) found very slight overlap between ratings and patient perceptions and concluded that they were measuring different aspects of the relationship. Rogers et al. (1967) found gross agreement between patient perceptions and rated conditions and clear divergence between the therapist's views of the relationship and views of both judges and patients.

The Question of Causation

Several investigators have attempted to determine whether it is the patient or the therapist who causes the therapeutic conditions.

Truax and Carkhuff (1966) reported a study in which 24 patients were seen by eight different therapists. Samples were selected in which each of the eight therapists had seen the same eight patients. Analysis of ratings indicated that different therapists tended to provide different levels of accurate empathy when interacting with the same patients ($p < .01$). Different patients did not tend to receive different levels of accurate empathy when interacting with the same therapists. Further analysis of the same data on measures of nonpossessive warmth and genuineness yielded similar findings. Different therapists tended to produce different levels of nonpossessive warmth ($p < .01$) and
genuineness \( (p < .01) \) with different patients. Different patients did not evoke different levels of nonpossessive warmth or genuineness when seen by different therapists.

In the study reported by Truax et al. (1966) with 40 outpatients at the Phipps Psychiatric Clinic, analysis of variance indicated significant differences between therapists in the level of conditions offered their respective sets of patients on accurate empathy \( (p < .001) \), genuineness \( (p < .001) \), and nonpossessive warmth \( (p < .05) \) \( (p = .102) \).

Part of the same study did suggest that patients as well as therapists affect the level of nonpossessive warmth, at least initially \( (p = 102) \). Two separate standard screening interviewers met with each of 40 patients prior to therapy. Results of analysis of variance for screening interviewer effects indicated significant differences between the interviewers on accurate empathy \( (p < .001) \) and genuineness \( (p < .01) \) but no differences on nonpossessive warmth. The investigators believed that the levels of accurate empathy and genuineness were under the immediate and direct control of the therapist whereas the level of nonpossessive warmth offered by the therapist is established over time but is initially influenced by the nature of the patient. Other studies support the hypothesis that the levels of therapeutic conditions offered throughout counseling are due to the counselor rather than the client (Hirshberg, Carkhuff, and Berenson, 1966; Banks, Berenson, and Carkhuff, 1966).
In contrast, Rogers et al. (1967) found certain initial patient factors to be significantly associated with the level of therapeutic conditions experienced by the patient. A relatively higher level of conditions (rated accurate empathy and patient perceived congruence) was generally related to the following initial patient factors: high socio-economic status, male sex, chronicity, high verbal intelligence, high verbal productivity in the second therapy interview, high mental health rating on the initial TAT, and a generally low level of manifest psychotic disturbance on the Wittenborn Psychiatric Rating Scales. The investigators suggested that initial patient factors and therapist attitudes in turn may constitute the primary factor in sustaining meaningful behavioral change. They suggest that as the therapy interaction progresses, general patient capacities, motives, and level of experiencing are so inextricably bound together with therapist attitudes that it seems most appropriate to conceive of therapy outcome as a complex function of their dynamic interaction (p. 309). In a related study, VanderVeen (1965) concluded that the main effects of both therapist and patient were significant for the rated genuineness and accurate empathy scales. The variation due to the therapist was consistently greater than that due to the patient. In addition, therapeutic behavior by one participant was associated with more therapeutic behavior by the other.
Discussion

Evidence concerning the question of causation is far from complete. It would appear that while both therapists and clients bring attitudes and characteristics into the relationship which affect the atmosphere and outcome of that relationship, therapist effects tend to be more significant. However, the view that conceives of therapy outcome as a complex function of the dynamic interaction between therapist and patient would seem most appropriate.

The Open Mind and Provision of Therapeutic Conditions

Evidence is becoming available which supports the theory that the open-minded individual is more likely to provide genuineness, nonpossessive warmth, and accurate empathy when participating in the counselor-client relationship.

Kemp administered Form E of the Rokeach Dogmatism Scale and Porter's test of Counselor Attitudes to 25 students before and after participating in a counselor practicum and to 25 students who did not participate in the practicum (Kemp, 1962). The 25 experimental cases were rated in actual counseling interviews on the same five categories of the Porter Scale: (1) evaluative or value setting, (2) interpretative or teaching responses, (3) understanding responses, (4) supportive responses, and (5) probing or diagnostic responses.
Members of the control group did not change significantly from pre to post testing. Counselors who were high on the Dogmatism Scale had fewer understanding and supportive responses on the Porter test.

In the experimental group, all participants changed in each of the five categories of the Porter test. The change was toward permissiveness and understanding. Those with open belief systems also made a comparable change in the actual counseling interview. However, there was a significant change for the closed belief group between the second administration of the Porter test and actual counseling ($p < .01$). The direction of change was toward more evaluative, interpretative, and probing or diagnostic responses. The apparent gains in the practicum as shown by the Porter test did not transfer to the counseling interview.

The investigator concluded that "the open minded, who could be expected to be more aware of their reactions to stimuli, who have less need to narrow and distort, and who normally consider ideas on their merits, are better integrated, experience less threat, have less anxiety, and are more permissive in their normal relationships. They respond this way and therefore are more understanding and supportive in their responses" (p. 156).

Stefflre, King, and Leafgren (1962) asked 40 National Defense Education Act (NDEA) Guidance Institute members to rate their 39 peers in order of effectiveness in the
counseling relationship. Consistency among judgment was remarkably high ($r = .96$). The top and bottom nine were designated as effective and ineffective counselors. The effective participants had higher academic performance, more appropriate Strong Vocational Inventory scores, and were less dogmatic as measured by the Rokeach scale. The effective counselors had a mean dogmatism score of 125.9 compared with a mean of 142.3 for the ineffective counselors.

In a similar study, Milliken and Paterson (1967) administered the Rokeach Dogmatism Scale to 30 NDEA Guidance and Counseling Institute enrollees. A Negro was employed as a coached counselee to be counseled by each trainee. Effectiveness of the enrollee's counseling was assessed by the coached client and the supervisor. The Mann-Whitney U was used to test differences in the ranks of dogmatism scores between good and poor counselors. The good counselors as defined by the coached client's global rating and the coached client's composite score scored insignificantly lower on the Rokeach Scale. The good counselors as defined by the supervisor's composite score were significantly lower on the Rokeach Scale ($p < .05$).

In a related study, Kemp (1964) administered the Rokeach Dogmatism Scale and the Zukerman Affect Adjective Check List for the measurement of anxiety to 150 college students ranging in age from 20-25. In addition, each participant was requested to estimate his own degree of open-mindedness.
and anxiety. Those with closed belief systems perceived themselves as more open minded. The open minded were able to perceive their anxiety more accurately than the closed minded.

Saltzman (1966) noted in his doctoral dissertation that the degree to which a person is initially perceived by others of being empathic, congruent, and unconditionally positive in his regard for others appears to be a function of his level of dogmatism. Using the Barrett-Lennard Inventory with 30 members of a NDEA Guidance and Counseling Institute, he found that persons who were low in dogmatism as measured by Rokeach Scale were viewed most positively by peers. However, those persons high in dogmatism were rated more positively than persons scoring in the middle group on the Rokeach Scale.

Discussion

From the definition of the highly dogmatic person, it would appear that he would be unlikely to provide high levels of therapeutic conditions. Limited research involving counselors suggests that highly dogmatic persons are less inclined naturally to provide therapeutic conditions and also find it more difficult to adapt their behavior toward providing higher levels of therapeutic conditions. Supervisors tended to describe trainees low in dogmatism as
better counselors. Fellow counselors viewed open minded peers as being more empathic, congruent, and unconditional in regard as measured by the Barrett-Lennard Relationship Inventory.

This chapter has presented research related to (1) essential ingredients of the therapeutic relationship, (2) ratings of therapeutic conditions, (3) perceived therapeutic conditions, (4) the relationship between rated and perceived conditions, (5) the question of causation, and (6) helper dogmatism.
CHAPTER III

PROCEDURES

The purpose of this study was to find support for the factors in the "helping person" that are hypothesized to bring about therapeutic results and to identify those persons who are most likely to bring about positive behavioral change.

This chapter describes procedures for selection of the student population and the helpers, the instruments used, and the procedures for collection and analysis of data.

Selection and Description of Community Helper Children

Any child in grades one through six referred to the school psychologists in the Franklin County, Ohio, public schools was tentatively eligible to participate in the Community Helper Program. The school psychologist then referred to the program coordinator (the author) those children who appeared to meet the qualifying criteria. The following criteria were used:

GRADE 3 AND UP

(1) Behavioral problem—manifestation of a behavioral disturbance in school (interpersonal difficulties, lack of self-confidence, withdrawal, defiance of
authority, aggressiveness, acting out, etc.). In some but perhaps not all cases there will be a suggestion of parent-child conflict.

(2) IQ of 80 or above.

(3) School achievement (one of the following):

(a) One grade year or more retarded in grade placement in comparison with Educational Age (Mental Age).

(b) One grade year or more retarded in achievement on a standardized achievement test (reading and/or arithmetic) in comparison with Educational Age (EA).

(c) Failing or marginal grades (D in one or more subjects).

GRADE 2 and BELOW

(1) Behavior problem (as above)

(2) IQ of 80 or above.

(3) School achievement (one of the following):

(a) One grade year or more retarded in grade placement in comparison with EA.

(b) Six months or more retarded in achievement on a standardized achievement test (reading and/or arithmetic) in comparison with EA.

(c) Failing or marginal grades (D in one or more subjects).

Children who were eligible for the Community Helper Program were randomly assigned whenever possible to either the experimental or control group. However, when a child was known to be receiving other treatment, he was assigned to the control group. Other treatment included remedial reading, speech therapy, or private psychological or psychiatric consultation.
Before being assigned to the Community Helper Program, each child was evaluated by an individual intelligence scale (Wechsler Intelligence Scale for Children or Stanford-Binet, L-M), the Gray Oral Reading Paragraphs, the Wide Range Achievement Test, a picture projective technique, and two behavior check lists completed by the teacher.

The student population for this study consisted of 35 of the 37 elementary students who made up the experimental cases in the Community Helper Program. Two were eliminated because of incomplete data. Two others were included for partial analysis even though outcome data were incomplete. There were 27 boys and eight girls ranging in ages from 7.1 to 12.0 with a mean age of 9.3 years. One child was in grade six, five were in grade five, 11 were in grade four, seven were in grade three, ten were in grade two, and one was in grade one. Thirteen had repeated one or more grades. Students came from 15 elementary schools in six different county school districts. The mean intelligence quotient for the subjects was 100.24 with a standard deviation of 11.16. Students tended to come from lower-middle-class families as measured by the North-Hatt Scale (North and Hatt, 1947).
Selection and Description of Community Helpers

Each experimental child was assigned to a Community Helper. A Community Helper was an adult volunteer who visited the school and who worked with the child twice each week in individual sessions. The helper was selected by the principal because of her apparent liking for children and willingness to participate. She was told that the main goal was to establish a "good" relationship between child and adult and was encouraged to use any activities which she liked in order to accomplish this goal. The program coordinator offered to help secure any needed materials and to meet with the helper periodically to discuss activities and progress. Each helper met with the coordinator from four to eight times during the treatment period. The coordinator attempted to provide a warm and non-threatening atmosphere so that the helper would feel free to "be herself" with the child. The number of helper-child interviews ranged from 14 to 25 with a mean of 22.1 (Appendix A).

Each helper was asked to complete a personal data sheet before being assigned to a child. Helpers were married women who had or were in the process of raising their own children. They ranged in age from 25 to 60 with a mean age of 38. Thirty-two of the 37 helpers fell between the ages of 29 and 45. Number of helpers' children ranged from one to five with a mean of 2.7. Two helpers had five
children while four had only one. Ages of helpers' children ranged from two to 39.

All but one of the helpers had graduated from high school. Twenty had begun college but only six had received BA or BS degrees. Mean level of educational attainment was 13.3 years. None had attended graduate school. One helper had a BA degree in elementary education. Another was attending college and working toward certification as a teacher while participating in the Community Helper Program.

None of the helpers were working full time at the beginning of the experimental program. However, most had been employed at one time and several were considering returning to work at a later date. Seventeen stated that their previous employment had been secretarial. Three had been nurses. Only one had ever been a teacher in the public schools.

It appeared that most of the husbands of the helpers were relatively well trained and successful in their work. Among them were one graduate physicist, nine graduate engineers, one public accountant, one minister, one air force pilot, and two high school teachers. They tended to be from upper-middle class homes as measured by the North-Hatt Scale (North and Hatt, 1947).

The personal data sheets revealed that helpers were active with hobbies and community activities. Favorite hobbies included sewing, sports, reading, music, bridge, art,
flower arranging, and knitting. All helpers with children in school belonged to their local PTA and most were holding or had held an executive position. Twenty-one indicated that they were active in their churches, often as teachers in the elementary classes. Six were 4-H or scout leaders at the beginning of the program. Six helpers belonged to the Child Conservation League and five belonged to Republican Women's Clubs.

In spite of their busy schedules, helpers stated that they were participating in the program because they had extra time and wanted to do something useful with it. Nineteen said that they enjoyed working with children. Six helpers mentioned that they had wanted to be teachers but had not been able. At least three still hoped to go back to college for teacher training. Seven felt that the experience of working with someone else's child might help them better understand their own children. In nearly every case, the helper seemed to be saying that it would make her feel good if she knew that she had helped someone.

Two helpers refused to complete the Rokeach Dogmatism Scale stating that the instrument was unnecessarily prying into their privacy. For the same reason, only 19 of the 37 helpers completed the Edwards Personal Preference Scale. The results of the Edwards Personal Preference Scale were not included in this study.
The Instruments

The three scales developed by Truax (1961, 1962a, 1962b) were used to obtain a measurement of the therapeutic conditions which were present in the helper-child relationship. Judges listened to tape-recorded segments of the helper-child interviews and rated them according to the instructions of the scale. Truax offers the following descriptions of his scales:

1. Tentative Scale for the Measurement of Therapist Self-Congruence (called genuineness) (Appendix B). The measurement of therapist genuineness from recorded psychotherapy sessions uses a seven-point scale descriptively specifying stages along a continuum. At its lowest level the scale includes such descriptions as "... there is explicit evidence of a very considerable discrepancy between his experiencing and his current verbalizations," and "... the therapist or counselor makes striking contradictions" in his statements ... or, the therapist may contradict the content ... with the voice qualities ... ." At higher values of the continuum, "there is neither implicit nor explicit evidence of defensiveness or the presence of a facade," and at the highest point there is an openness of experiences and feelings by the therapist of all types—both pleasant and hurtful—without traces of defensiveness or retreat into professionalism . . . " (Truax, 1962a).

2. Tentative Scale for the Measurement of Unconditional Positive Regard (called nonpossessive warmth) (Appendix C). The measurement of nonpossessive warmth specifies a five-step continuum involving at the lower range such helping behaviors as (he) acts in such a way as to make himself the locus of evaluation . . . (he) may be telling the patient what would be "best" for him, or may be in other ways actively trying to control his behavior, or, the therapist "responds mechanically to the client and thus indicates little positive warmth . . . or ignores the patient where an unconditionally warm response would be expected—complete passivity that communicates a lack of warmth." At very high values (he) clearly communicates a very deep interest and concern for the
welfare of the patient. Attempts to dominate or control the patient are for the most part absent . . . except that it is important that he (the patient) be more mature . . . or that the therapeutic person himself is accepted and liked, "or at the highest level . . . the patient is free to be himself even if this means that he is temporarily regressing, being defensive, or even disliking or rejecting the therapist himself" (Truax, 1962b).

3. A Scale for the Measurement of Accurate Empathy (Appendix D). The accurate empathy scale defines a nine-step continuum which specifies at its lower values such behaviors as "he seems completely unaware of even the most conspicuous of the patient's feelings. His responses are not appropriate to the mood and content of the client's statements and there is no determinable quality of empathy, hence no accuracy whatsoever." Whereas, at intermediate levels of the continuum he often responds accurately to more exposed feelings. He also displays concern for more hidden feelings which he seems to sense must be present, though he does not understand their nature." Or, "he shows awareness of many feelings and experiences which are not so evident . . . but in these he tends to be somewhat inaccurate in his understanding." At the higher levels of the continuum of accurate empathy, the therapist "shows awareness of the precise intensity of most underlying emotions . . . his responses move only slightly beyond the area of the client's own awareness, so that feelings may be present which are not recognized by the client's present, acknowledged feeling." He moves into feelings and experiences that are only hinted at . . . and does so with sensitivity and accuracy. (He) offers additions to the patient's understanding so that not only are underlying emotions pointed to, but they are specifically talked about" (Truax, 1961).

The scales were designed for use with live observations or tape recordings of counseling or therapy interviews. The author states that they may be used on typescripts with only slight loss of reliability. They have been used with various lengths of time samples and in both individual and group interaction.
Table 1 (p. 27) presents the interrater reliability coefficients from seventeen studies involving a variety of therapist and patient populations. Correlations range from .40 to .89 for all three variables. Rogers et al. (1967) reported lower interrater reliabilities than any of the studies reported by Truax (accurate empathy, .36, nonpossessional warmth, .08).

Concerning validity, Truax and Carkhuff (1966) state that "We know from the evidence . . . that these scales are significantly related to a variety of client therapeutic outcomes" (p. 44). They believe that the scales measure central therapeutic ingredients which should be central aspects of training and practice in counseling and psychotherapy.

**Relationship inventory**

The relationship inventory was used to determine how the child and the helper viewed the helper in the relationship (Appendix E). Existing in two parallel forms, the inventory consisted of 24 true-false items. It measured level of regard, accurate empathy, genuineness, and conditionality of regard. There were six questions for each of the four variables. The child was asked to respond to such items as "She likes seeing me!" The helper's corresponding item read, "I like seeing the child."

Barrett-Lennard developed a 92 item relationship inventory which included these four variables and one other,
"willingness to be known" (Barrett-Lennard, 1962). His questionnaire consisted of items that could be responded to on a six-point continuum from strong agreement (+3) to strong disagreement (-3). The items representing each variable (level of regard, empathy, congruence, and unconventionality of regard, willingness to be known) are dispersed throughout the inventory so as to obtain maximum independence of answers. Content validation was carried out by the use of ratings by five judges and by a formal item analysis.

The instrument used in this study was a greatly shortened and simplified version of the Barrett-Lennard Scale. Wording and sentence structure were simplified so that it could be understood by children as young as seven years. The six-point continuum was changed to the true-false items. The number of items was reduced from 92 to 24 and the "willingness to be known" factor was eliminated completely since it did not relate to the present study.

In short, an attempt was made to preserve the content validity established by Barrett-Lennard while making the inventory applicable to small children.

The Rokeach Dogmatism Scale (Form E)

The Dogmatism Scale was developed by Rokeach (Appendix F). Its theoretical framework and development have been described in his book The Open and Closed Mind (1960).
The present scale is the latest of five revisions which have attempted to continue refinements in theoretical formulations and to increase test reliability. Form E contains 40 items. In the form used for this study, 20 additional neutral items were included making a total of 60. Each item is responded to on a six-point continuum ranging from strong agreement (+3) to strong disagreement (-3). The higher the score, the more "dogmatic" or "closed-minded" the person taking the inventory.

The raw score for each item is found by adding four points to the actual value of each response. A response of -3 thus becomes +1, as minus scores are transformed into positive scores.

Table 2 reports a series of studies which indicate total means, standard deviations, and reliabilities for selected groups of subjects on the Dogmatism Scale.

Procedures for Collection of Data

Index of therapeutic success

A major difficulty in a study such as this is that of finding a valid measure of positive behavioral change. Four separate criteria were available from the Community Helper Program and used in the present study. In each instance, gain scores derived from evaluations before and after treatment for each child were changed to standard scores and then to T-scores. In this way, four indexes of success
TABLE 2
MEANS, STANDARD DEVIATIONS, AND RELIABILITY COEFFICIENTS FOR THE ROKEACH DOGMATISM SCALE (FORM E)

<table>
<thead>
<tr>
<th>Number of Cases</th>
<th>Reliability</th>
<th>Mean</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>English Colleges 80</td>
<td>.81</td>
<td>152.8</td>
<td>26.2</td>
</tr>
<tr>
<td>English Workers 60</td>
<td>.78</td>
<td>175.8</td>
<td>26.2</td>
</tr>
<tr>
<td>Ohio State U. I 22</td>
<td>.85</td>
<td>142.6</td>
<td>27.6</td>
</tr>
<tr>
<td>Ohio State U. II 28</td>
<td>.74</td>
<td>143.8</td>
<td>22.1</td>
</tr>
<tr>
<td>Ohio State U. III 21</td>
<td>.74</td>
<td>142.6</td>
<td>23.3</td>
</tr>
<tr>
<td>Ohio State U. IV 29</td>
<td>.68</td>
<td>141.5</td>
<td>27.8</td>
</tr>
<tr>
<td>Ohio State U. V 58</td>
<td>.71</td>
<td>141.3</td>
<td>28.2</td>
</tr>
<tr>
<td>V.A. Domiciliary 80</td>
<td>-</td>
<td>183.2</td>
<td>26.6</td>
</tr>
<tr>
<td>OSU, NDEA Guidance and Counseling Institute (Saltzman, 1966) 30</td>
<td>-</td>
<td>130.6</td>
<td>-</td>
</tr>
<tr>
<td>American Teachers in summer school (Rabkin, 1966) 107</td>
<td>-</td>
<td>132.2</td>
<td>22.5</td>
</tr>
</tbody>
</table>

were established for the purpose of evaluating the validity of measurement variables. In addition, a total outcome index was derived by combining the four index measures into one.

Intelligence.—Each child was administered a Stanford-Binet Intelligence Scale or Wechsler Intelligence Scale for Children at the beginning and at the end of the treatment
program. If the referring psychologist had already administered one of the two instruments within a month of referral, it was not readministered. In cases where the referring psychologist had not administered either scale, a Wechsler Intelligence Scale for Children was given before beginning the program. The instrument administered before treatment was again administered at the conclusion. Gain scores measured in intelligence quotient points were changed to T scores.

Achievement.—The available evidence for change in achievement was combined into one index. Reading and arithmetic scores on the Wide Range Achievement Test, scores on Form A of the Gray Oral Reading Test, and report card grades were available for each child before and after treatment. The two reading scores were combined to form a mean reading score. This score was then averaged with the arithmetic score, adjusted for time in treatment, and changed to a T-score. Gain in grade point average was also transformed into a T-score and averaged with measured achievement T-scores to form the total achievement index.

Behavior.—For the purpose of evaluating behavior change, the teacher of the child was asked to complete two behavior check lists before and after treatment. The Quay-Peterson Behavior Problem Checklist consists of 55 samples of problem behaviors for the teacher to rate (Peterson, 1961, and Quay and Quay, 1965). Each item is to be rated 0— not a problem, 1—a mild problem, 2—a severe problem. The higher
total score denotes the more serious behavior disturbance. Each item purports to measure disturbance in one of three behavioral classifications: (1) conduct disorders, unsocialized aggression, or psychopathy; (2) personality problem or neuroticism, and (3) inadequacy-immaturity. Total gain scores (gains represented by an actual reduction in number of items checked by teacher from pre to post testing) were transformed into T-scores.

An Interpersonal Adjective Check List was employed to assess interpersonal behavior. Consisting of 88 items, it was administered and scored in the same fashion as the behavior problem check list. Adapted from diagnostic categories developed by Leary (1956) and LaForge and Suczek (1955), the interpersonal check list classifies behavior into the following areas: managerial-autocratic, responsible-hyper-normal, cooperative-overconventional, docile-dependent, self-effacing-masochistic, rebellious-distrustful, aggressive-sadistic, competitive-narcissistic, and withdrawn-outgoing. Total gain scores (reduction in number of times checked) from this instrument were also transformed into T-scores. The final index of behavior change consisted of the mean of the T-scores for the Quay-Peterson Behavior Problem Check List and the Interpersonal Adjective Check List.

Motivation.—Thirteen stimulus pictures, selected from the Thematic Apperception Test and the Michigan Picture Story Test, were administered to each child and scored according to Arnold's Story Sequence Analysis (1962) before and after
the treatment program. This method of analysis attempts to
discover positive and negative motivation. The author cites
evidence that positive motivation is found among high
achieving elementary, secondary, and college students, among
effective teachers, competent executives, and well adjusted
Navy men. Negative motivation is found among low achievers,
14). The four scoring categories include (1) Achievement,
success, happiness, active effort (or lack of it); (2) Right
and wrong; (3) Human Relationships; and (4) Reaction to
Adversity. Total gain scores on this were transformed into
T-scores to form the motivation outcome index.

Total index.—A total outcome index was obtained by
averaging the T-scores from the four measures. In order to
do this, arbitrary decisions as to weighting were made.
Since it was expected that the therapeutic relationship would
be most influential in the areas of behavior and achievement,
T-scores on these two variables were given double weighting;
gains in intelligence and motivation were given single weight-
ing. The mean of the resulting six T-scores was then used as
the index of therapeutic success for each child.

Selection of interview segments
to be rated

For each experimental case, audio tape recordings were
made of two consecutive interviews early in the community
helper program and two consecutive interviews late in the
program. Whenever possible, the third, fourth, 23rd and 24th
interviews were taped. In cases which did not reach 25 interviews, tapes were made of the second from last and third from last interviews. Scheduling conflicts caused some variation. Whenever possible, the second and fourth tapes were used for analysis. The first and third tapes were included in order to help the participants get used to being taped. Because of the inaudibility of several tapes, this procedure could not always be followed. The number of interviews and tapes used for analysis and the total number of interviews for each case before re-evaluation took place can be found in Appendix A. Number of interviews between early and late tapes ranged from ten to 21 with a mean of 17.3.

For each of the selected tapes, three three-minute segments were chosen for analysis. Selection of segments was limited both by the audibility of the tapes and the amount of relevant interaction between helper and child. Tapes were considered audible if the secretary was able to transcribe them on to typescript. However, some segments contained no personal interaction between helper and child. For example, a segment might contain continuous reading by the helper or the child or some other completely academic task such as spelling where no relevant interaction was detectable via audio tapes.

Whenever possible, the 0-3, 9-12, and 18-21 minute segments were used for analysis. When the selected segment was judged unsuitable because of lack of interaction between
helper and child, the investigator took the three minute segment beginning with the first relevant interaction following the designated starting point. Some segments contained much more interaction than others.

**Rating procedure.** Before rating took place, the 70 tapes were placed in random order. The selected segments were then re-taped in the proper order for ease of rating. The final product was 630 minutes of taped material (210 three-minute segments).

The rater did not know the helper or the child; whether it was an early or later interview; or whether the case had been successful or unsuccessful. The rater was informed that only the tapes and not the segments within an interview had been randomized. The rating form used by the rater contained three identical scales, one for each segment of that particular interview. The rater was asked to mark the proper stage for each segment. All raters completed Scale I (Non-possessive Warmth) before beginning Scale II (Accurate Empathy). Scale III (Genuineness) followed the completion of Scale II. The same tape sequence was used for each variable. Each recording was accompanied by transcript for each of listening. Upon completion of the training period, raters were allowed to work alone at their own convenience. When in doubt, they were instructed to re-play all or any part of a segment.

**Raters and training procedure**—The three raters were advanced graduate students at Ohio State University in the area of School Psychology. Judges one and three were male;
respective ages were 25, 26, and 29. All three had taken an introductory course in counseling and had had considerable training in interviewing. In addition, judge number three had participated in a counseling practicum. By training and experience, judge three should have been slightly more sophisticated than judges one and two.

The training procedure was identical for each of the three rating scales. Each rater was given a copy of the scale with a full description of each stage. Later, the investigator met jointly with the raters to discuss their conceptions of what they would be rating. They were encouraged to ask questions and to try to anticipate difficulties. They were not being instructed. The purpose was to resolve conceptual differences on the part of the raters.

After raters appeared to have resolved conceptual differences, they rated five sample segments with the investigator. After hearing a segment, each of the four raters gave his rating and tried to explain how he had arrived at that particular score. When differences appeared, they were examined for possible differences in theoretical orientation. This practice also gave raters a chance to establish a common baseline. At the end of five segments, raters felt that they were trying to measure the same thing and that they agreed upon the meaning of each stage.

Participant perceptions

Each helper completed the relationship inventory near the conclusion of the experimental treatment program. Each
child also completed the parallel form of the inventory at the time of re-evaluation. The examiner read the instructions and statements to the child to be sure that the child fully understood the task.

**Helper Dogmatism**

Each helper was asked to complete Form E of the Rokeach Dogmatism Scale. In most cases, this was completed by the fifteenth interview.

**Procedures for Analysis of Data**

**Reliability of ratings on interviews**

In order to establish reliability of the rating procedures, Pearson product-moment intercorrelations were computed for each combination of judges. A mean correlation for all three judges was determined by changing the correlations to Fisher's Z-scores, averaging them, and then changing the resulting Z-score back to a correlation coefficient.

**Mean, range and variability of ratings**

In order to determine the characteristics of rated conditions, means, range and variability of ratings were compiled and compared with corresponding measures from the Wisconsin Study (Rogers et al., 1967).

**Testing the hypotheses**

This study is designed to investigate the validity of the following hypotheses:

1. As judged by trained raters,
a. There will be a significant positive correlation between positive behavioral change in the child and level of genuineness in the helper.

b. There will be a significant positive correlation between positive behavioral change in the child and the level of nonpossessive warmth provided by the helper.

c. There will be a significant positive correlation between positive behavior change in the child and the level of empathic understanding provided by the helper.

Where inter-judge reliability coefficients were high enough to merit combining rater scores, Pearson product-moment correlations were computed between each of the measured conditions (accurate empathy-early, accurate empathy-late) and each of the five outcome indexes (intelligence, achievement, behavior, motivation, total index).

2. As measured by the relationship inventory,

a. There will be a significant positive correlation between positive behavioral change in the child and the child's relationship inventory total score.

b. There will be a significant positive correlation between positive behavior change in the child and the helper's relationship inventory total score.

Pearson product-moment intercorrelations were used to examine the relationship between the perceptions of the child and those of the helper and between rated conditions and the perceptions of the child and of the helper. In order to establish the relationship between helper and child perceptions and case outcome, Pearson product-moment correlations were computed between total scores of the relationship inventory and each of the five outcomes indexes.
3. As measured by Form E of the Dogmatism Scale, there will be a significant negative correlation between positive behavioral change in the child and level of dogmatism, on the part of the helper.

Pearson product-moment intercorrelations were used to examine the relationship between helper dogmatism and each of the rated conditions and between helper dogmatism and helper and child perceptions of the relationship. In order to establish the relationship between helper dogmatism and case outcome, Pearson product-moment correlations were computed between dogmatism scores and each of the five outcome indexes.

**Multiple correlations**

In order to account for the greatest possible amount of outcome variance, multiple correlations were determined using late-rated non-possessive warmth, late-rated accurate empathy, total score on the child relationship inventory, total score on the helper relationship inventory, and the dogmatism score as the five independent variables. The five dependent variables were each of the five outcome indexes.

The present chapter has described the procedures used in the study. It has indicated the nature of the student and helper populations, the nature of the instruments, and the procedures for collection and analysis of data.
CHAPTER IV

STATISTICAL ANALYSIS AND INTERPRETATION OF DATA

Chapter IV presents the statistical analysis and interpretation of the data. Preliminary topics in the chapter include the reliability of ratings of interviews (p. 66); means, range, and variability of ratings (p. 68); stability of rated conditions (p. 73); characteristics of relationship inventory scores (p. 75); characteristics of dogmatism scores (p. 76); and intercorrelations among predictor variables (p. 77). The hypotheses are then presented, together with the data and discussion concerning each hypothesis (pp. 82 to 85). Results of multiple correlations are presented (p. 85).

Reliability of Ratings

The score given to an interview on any one of the three rated variables (nonpossessive warmth, accurate empathy and genuineness) was to consist of the total score from nine rated segments (three judges each rating three segments). Before combining scores, however, it was necessary to determine the interjudge reliability for each of the three scales. Pearson product-moment correlations were computed from the sums of the ratings for each interview of each judge. Hence,
the basic unit of analysis became the total interview rather than a single segment. The correlations are shown in Table 3.

Reliability coefficients between raters one and three in nonpossessive warmth were very high (r=.94), indicating that the judges came very close to perceiving and evaluating the interviews in the same manner. Judge two also agreed on an acceptable level (r=.56, \( r = .54 \)) with each of the other two raters. Coefficients were changed to Fisher's Z scores, averaged, and changed back to correlation coefficients. The mean correlation coefficient for rated N.P.W was .75. The judgments of the three raters were pooled for all further analyses.

**TABLE 3**

<table>
<thead>
<tr>
<th>Variables</th>
<th>( r_{12} )</th>
<th>( r_{23} )</th>
<th>( r_{13} )</th>
<th>Mean ( r_b )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonpossessive Warmth</td>
<td>.56</td>
<td>.54</td>
<td>.94</td>
<td>.75</td>
</tr>
<tr>
<td>Accurate Empathy</td>
<td>.53</td>
<td>.59</td>
<td>.92</td>
<td>.74</td>
</tr>
<tr>
<td>Genuineness</td>
<td>.18</td>
<td>.37</td>
<td>.04</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\)Correlation coefficients were subscripts 1, 2, and 3 refer to judges one, two, and three, as stated earlier.

\(^b\)Averages computed using Fisher's Z scores.
Reliability coefficients between judges on accurate empathy were similar to those for nonpossessive warmth. Judges one and three agreed highly ($r = .92$) while judge two agreed but to a lesser extent ($r = .53$, $r = .59$) with the other two. Coefficients were changed to Fisher's Z scores, averaged, and changed back to correlation coefficients. The mean correlation coefficient for rated accurate empathy was .74. Again, the judgments of the three raters were pooled for further analyses.

Reliability coefficients on the genuineness scale revealed that raters were unable to agree on this variable. Only the correlation between judgments by raters two and three approaches a level acceptable for pooling ($r = .37$). Raters commented that this variable was difficult to measure because helpers seldom expressed their own feelings. Because judges could not agree on this variable, it was excluded from further analysis except in the discussion of means, range and variability of ratings.

**Means, Range and Variability of Ratings**

It is important to note the range and distribution of ratings for each of the three scales. To make them more meaningful, they have been compared with scores reported from the Wisconsin Study (Rogers et al., 1967).

Figure 1 presents the percentage that each stage was used in rating nonpossessive warmth for the Community Helper
A description of each scale stage is contained in Appendix B.

C - Community Helper Program \(N=630\) Mean 2.6.

W - Wisconsin Study \(N=1011\).

Fig. 1.—Per cent of ratings assigned to each stage\(^a\) of the nonpossessive warmth scale in the community helper program and the Wisconsin study.
Programs and the Wisconsin Study. Explanation of each stage can be found in Appendix C. The higher stages represent higher levels of rated nonpossessive warmth. In the Wisconsin Study, the bulk of the rated segments fell into the top three stages. Stage three was most often rated, accounting for 45 per cent of the segments. Only 8 per cent of the segments fell into stages one and two. Stage three was also rated most often in the present study. However, 43 per cent of the ratings fell at lower scale points with only 17 per cent at levels four and five. In the present study, the mean rating scores for both early and late interviews was 2.6. This would suggest that, as a group, community helpers did not change their level of nonpossessive warmth from early to late interviews. As judged by raters, community helpers tend to provide relatively low levels of nonpossessive warmth when compared with the trained therapists who made up the Wisconsin Study.

The Wisconsin Study and the present study also used the same scale to measure accurate empathy. Figure 2 presents the percentage that each stage was used in both studies. Appendix D contains the detailed explanation of each stage. The higher stages indicate higher levels of accurate empathy. Comparison of ratings suggests a marked difference between studies. In the Wisconsin Study, scale ratings cover nearly the full range of the scale with a midpoint at the natural midpoint of the scale, stage five.
A description of each scale stage is contained in Appendix D.

C - Community Helper Program  N=630  Mean 3.3.
W - Wisconsin Study N=1011.

Fig. 2.--Per cent of ratings assigned to each stage\textsuperscript{a} of the accurate empathy scale in the Community Helper Program and in the Wisconsin Study.
They suggest that the therapists usually had a fairly accurate awareness of the patient's feelings. In contrast, 79 per cent of the ratings fell below the midpoint in the present study. Stages eight and nine were never used. The mean score of 3.3 in the present study suggests that community helpers were providing relatively low levels of rated accurate empathy when compared with therapists in the Wisconsin Study. However, other studies using trained therapists have also reported relatively low levels of rated accurate empathy (Bergin and Solomon, 1963; Melloh, 1964).

Even though judges in the present study were unable to differentiate between helpers on the basis of the genuineness scale, they did agree that helpers, as a group, were relatively low on this variable. Using the seven point scale contained in Appendix B, judges rated 77 per cent of the segments as neutral or ungenuine. The Wisconsin Study used a five point scale. In sharp contrast, the majority of the ratings for therapists fell at stage five. No therapist was judged as phony or defensive.

The data suggests that, on the average, Community Helpers were providing lower levels of the three therapeutic conditions than were the trained psychotherapists in the Wisconsin Study. It would appear that Community Helpers and therapists were most alike in the provision of nonpossessive warmth and least alike in level of genuineness. In other
words, Community Helpers were more likely to provide high levels of nonpossessive warmth than high levels of accurate empathy or genuineness.

It is unlikely that the rated differences between therapists and Community Helpers are a result of a different orientation on the part of raters although this may be a factor. Nor is it surprising that trained client-centered therapists are rated higher than untrained volunteers in the variables nonpossessive warmth, accurate empathy, and genuineness. Implications of these findings will depend upon the relevancy of the conditions in therapeutic success with children.

**Stability of Rated Conditions**

Table 4 presents the correlations between rated conditions from early to late interviews. Correlation between nonpossessive warmth from the early to late interview is low and insignificant \((r=.12)\). The relatively low correlation between nonpossessive warmth from an early to late interview suggests either that this condition has not yet stabilized by the third or fourth interview or that community helpers are not consistent from one interview to another in the level of nonpossessive warmth which they provide. Truax (1966) has suggested that both client and helper contribute to the initial level of nonpossessive warmth offered while the helper largely determines later levels. At any rate, it would
appear that the fourth interview is too early to predict the level of nonpossessive warmth which will be provided in later interviews.

### TABLE 4
PEARSON PRODUCT-MOMENT CORRELATIONS BETWEEN RATED CONDITIONS—EARLY AND LATE
N=35

<table>
<thead>
<tr>
<th>Variables</th>
<th>Nonpossessive Warmth (Late)</th>
<th>Accurate Empathy (Late)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nonpossessive Warmth (Early)</td>
<td>.12</td>
<td></td>
</tr>
<tr>
<td>Accurate Empathy (Early)</td>
<td></td>
<td>.35</td>
</tr>
</tbody>
</table>

Correlation between accurate empathy from the early to late interview is somewhat higher but still accounts for a relatively small amount of variability (r=.35). Rogers et al. (1967) reported that rated accurate empathy became relatively consistent at about the eighth interview and remained so through the 15th interview. Evidence from the present study could be interpreted as supporting both the Truax and Rogers hypotheses. Truax would suggest that early accurate empathy correlates more highly with late accurate empathy than would early nonpossessive warmth with late nonpossessive warmth because it is more under the direct control of the helping person. However, the correlations are still low enough to suggest that stabilization has not yet been completed for either variable.
Characteristics of Relationship Inventory Scores

The relationship inventory used for this study is a crude instrument patterned after the Barrett-Lennard Relationship Inventory (Appendix E). This instrument was employed in an attempt to determine whether or not helper or child perceptions of the relationship are related to case outcome. Data evolved from this instrument should be interpreted conservatively.

The inventory consists of 24 true-false items purporting to measure level of regard, conditionality of regard, accurate empathy, and genuineness. Each subscale score consists of the total positive responses from the six items of that subscale. Parallel forms were administered to helper and child.

Examination of the completed inventories revealed characteristics which should be kept in mind when interpreting the results. Helpers almost unanimously gave themselves perfect scores on the level of regard subscale, causing a highly skewed distribution of scores. Several other individual items did not distinguish between helpers. Finally, inventory administrators were not certain that all children understood all of the items.

In spite of these limitations, several tentative conclusions can be made from the relationship inventories. Helpers tend to perceive the relationship more positively than do the children. This is consistent with evidence
reported by Barrett-Lennard (1962) and Rogers et al. (1967). Only in accurate empathy did helpers rate themselves lower than children rated them. Both helpers and children rated helpers lowest in conditionality of regard (Appendix G).

Intercorrelations among subscales of the child relationship inventory tend to be high and positive. All subscales also correlate highly with the total score. In contrast, helper subscale scores do not appear to be highly related and correlate lower with the total score. In other words, the data suggests the relationship inventory total score represents a single global variable for children while helpers tend to differentiate between the four conditions (Appendix G). Because of the questionable reliability of individual subscales, only total scores on the relationship inventory were used for further analysis.

Characteristics of Dogmatism Scores

Helper scores on the Rokeach Dogmatism Scale, Form E ranged from 104 to 200 with a mean of 148.28 and a standard deviation of 19.99. There are no norms involving persons similar to the participants in the Community Helper Program. Comparison of the mean score for community helpers with mean scores for other groups would suggest that community helpers, as a group, score higher on the dogmatism score than do students in counselor training, American teachers, and American university students (Table 2, p. 56).
Intercorrelations Among Predictor Variables

Five measures were selected in order to further investigate the intercorrelations among predictor variables. Late-rated nonpossessive warmth, late-rated accurate empathy, the child relationship inventory total score, and the helper relationship inventory total score were selected since all four measures were taken near the end of treatment and should reflect the same interpersonal relationship. The Rokeach Scale was administered near the middle of the program but should be independent of time in the treatment program. These intercorrelations are found in Table 5.

Correlation between rated conditions

The high correlation between nonpossessive warmth and accurate empathy when rated in the same interview ($r = .70$) allows several possible interpretations. It may be that helpers who provide high levels of one condition are likely to provide high levels of the other, thus suggesting that both variables are related to an underlying basic attitude of the helper. It may also be that the two scales are not independent measures of the two different concepts, or at least not as seen by the present raters. Since the same judges rated all three scales, they may have been influenced by earlier ratings.

These findings are similar to those of Halkides (1958) who used similar scales. The rating procedure was
### TABLE 5

**PEARSON PRODUCT-MOMENT INTERCORRELATIONS AMONG FIVE PREDICTOR VARIABLES**

*N=35*

<table>
<thead>
<tr>
<th>Variables</th>
<th>Late Rated Accurate Empathy</th>
<th>Child Relationship Inventory Total Score</th>
<th>Helper Relationship Inventory Total Score</th>
<th>Helper Dogmatism Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Late-Rated Nonpossessive Warmth</td>
<td>.70**</td>
<td>.32*</td>
<td>.26</td>
<td>-.19</td>
</tr>
<tr>
<td>Late-Rated Accurate Empathy</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Child Relationship Inventory Total Score</td>
<td></td>
<td>.62**</td>
<td></td>
<td>.07</td>
</tr>
<tr>
<td>Helper Relationship Inventory Total Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Helper Dogmatism Score</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05 (one tailed test)

**p < .01 (one tailed test)**
also somewhat the same. Using the same three relatively sophisticated raters for all three scales, Halkides found an even higher correlation between accurate empathy and non-possessive warmth \((r=.89)\). Studies reported by Truax and Carkhuff (1966) indicate generally positive but lower correlations when using separate relatively unsophisticated judges for each scale. Rogers et al. (1967), using different undergraduate raters for each scale, obtained insignificant segment-by-segment Pearson intercorrelations among the three conditions scales.

From this study and those reported, several hypotheses can be suggested concerning the relationship between rated variables. Evidently more sophisticated judges tend to base their judgments on clinical inferences, causing the correlations between variables to rise. In addition, use of the same judges for more than one scale tends to increase the correlation between nonpossessive warmth and accurate empathy suggesting a halo effect. When relatively unsophisticated judges used for a single scale achieve a relatively high level of interjudge reliability, there tends to be a moderate positive interrelationship among the three conditions.

**Correlation between helper and child relationship inventories**

The data from Table 5 suggests that helper and child are in general agreement as to the global level of therapeutic
conditions \((r = .62)\). This is in contrast with previous studies. Rank-order intercorrelations among subscales and total scores of patient and therapist relationship inventories in the Wisconsin Study (Rogers et al., 1967) were generally insignificant and tended toward the negative direction. Barrett-Lennard noted a positive but insignificant trend between therapist and patient relationship inventory scores. Spotts (1962), however, using the Wisconsin data, reported that in the successful cases, there was a significant positive correlation between patient and therapist evaluations of the relationship, while in less successful cases therapist and patient assessments are more divergent, correlating negatively with one another.

**Correlation between rated and perceived conditions**

As evidenced by Table 5, the only other significant correlation among measured conditions is between the child relationship inventory total score and late-rated nonpossessive warmth \((r = .32)\). This, however, is considerably less than between rated conditions \((r = .70)\) and between perceived conditions \((r = .62)\). It would appear that the helper relationship inventory total score tends to be positively related to late-rated nonpossessive warmth \((r = .26)\) but the correlation does not reach the .05 level of significance. Neither the helpers' nor the children's relationship inventory total scores appear to be related to late-rated accurate empathy.
Rogers et al. (1967), working with hospitalized schizophrenic adults, reported that patients' general evaluations could be predicted with considerable accuracy from both judges accurate empathy ratings and from reliable nonpossessive warmth ratings but that therapists' relationship inventory evaluations and judges' ratings were either negatively associated or unrelated. Truax (1966) intercorrelated patient Relationship Inventory scores with rated conditions and reported that the Barrett-Lennard Relationship Inventory appears to be measuring different aspects of the therapy process.

As in the earlier studies, the data suggests that judges' ratings of conditions in an interpersonal relationship are measuring somewhat different aspects of that relationship than are child or helper perceptions. One might tentatively conclude, however, that the relationship inventory used in this study, especially when completed by children, is related to rated nonpossessive warmth.

Correlation between dogmatism scores and conditions measures

A negative correlation between helper dogmatism and measured therapeutic conditions would have indicated that more dogmatic persons were less likely to provide these conditions than less dogmatic persons. Although four of the five correlations with outcome indexes were in the negative direction, none were large enough to suggest a negative relationship (Table 5).
Intercorrelations between predictor variables and case outcome

Table 6 presents the intercorrelations between each of five predictor variables (late-rated nonpossessive warmth, late-rated accurate empathy, child relationship inventory total score, helper relationship inventory total score, helper Opinionaire score) and each of five outcome indexes (change in intelligence, achievement, number of behavior problems, motivation, and combined index). The hypotheses of this study are stated and discussed in relation to the data presented in this table.

Hypothesis la. As judged by trained raters, there will be a significant positive correlation between positive behavioral change in the child and the level of genuineness in the helper.

Because of lack of agreement between raters, a reliable measure of helper genuineness was not available and could not be included for analysis.

Hypothesis lb. As judged by trained raters, there will be a significant positive correlation between positive behavioral change in the child and the level of nonpossessive warmth provided by the helper.

Rated nonpossessive warmth is positively related to change in achievement ($p < .01$), reduction in behavior problems ($p < .05$), and to the combined outcome index ($p < .01$). It does not appear to be related to change in intelligence.
<table>
<thead>
<tr>
<th>Predictors</th>
<th>Intelligence</th>
<th>Achievement</th>
<th>Classroom Behavior</th>
<th>Achievement Motivation</th>
<th>Combined Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Late-Rated Nonpossessive Warmth</td>
<td>-.04</td>
<td>.41**</td>
<td>.32*</td>
<td>-.14</td>
<td>.40**</td>
</tr>
<tr>
<td>Late-Rated Accurate Empathy</td>
<td>.11</td>
<td>.37*</td>
<td>.16</td>
<td>-.14</td>
<td>.32*</td>
</tr>
<tr>
<td>Child Relationship Inventory Total Score</td>
<td>.23</td>
<td>-.01</td>
<td>.33*</td>
<td>.19</td>
<td>.41**</td>
</tr>
<tr>
<td>Helper Relationship Inventory Total Score</td>
<td>.23</td>
<td>.22</td>
<td>.07</td>
<td>.09</td>
<td>.29</td>
</tr>
<tr>
<td>Helper Opinionaire Score</td>
<td>-.09</td>
<td>-.18</td>
<td>-.18</td>
<td>.31</td>
<td>-.14</td>
</tr>
</tbody>
</table>

*p < .05 (one tailed test)  
**p < .01 (one tailed test)
or change in motivation. This data provides considerable support for hypothesis 1b.

**Hypothesis lc.** As judged by trained raters, there will be a significant positive correlation between positive behavioral change in the child and the level of empathic understanding provided by the helper.

Rated accurate empathy is positively related to change in achievement \((p < .05)\) and to the combined outcome index \((p < .05)\). The relationship does not extend to reduction in behavior problems as did nonpossessive warmth. Accurate empathy does not appear to be related to change in intelligence or change in motivation. The data might be interpreted as giving rather limited support for hypothesis lc.

**Hypothesis 2a.** There will be a significant positive correlation between positive behavioral change in the child and the child relationship inventory total score.

The child relationship inventory total score is positively related to reduction in behavior problems \((p < .05)\) and to the combined outcome index \((p < .01)\). It does not appear to be related to change in intelligence, change in achievement, or change in motivation. The data provides limited support for hypothesis 2a.

**Hypothesis 2b.** There will be a significant positive correlation between positive behavioral change in the child and the helper relationship inventory total score.
The helper relationship inventory total score correlates positively with each outcome index but only the correlation with the combined index nears the .05 level of significance. The data does not support hypothesis 2b.

**Hypothesis 3.** As measured by Form E of the Dogmatism Scale, there will be a significant negative correlation between positive behavioral change in the child and dogmatism on the part of the helper.

The dogmatism score correlates negatively with four of the five outcome indexes. None, however, approach a significant level. The data provides no support for hypothesis 3.

**Results of Multiple Correlations**

In order to determine the largest possible amount of outcome variance which could be accounted for by the predictor variables, multiple correlations were computed for each of the five outcome indexes. The five predictor variables included late-rated nonpossessive warmth, late-rated accurate empathy, the child relationship inventory total score, the helper relationship inventory total score, and the helper dogmatism score. The results of these multiple correlations are contained in Table 7.

None of the multiple correlations reach the .05 level of significance although this level was approached in the correlation for the combined index. Combining the predictor variables does not greatly increase the amount of variance
TABLE 7
MULTIPLE CORRELATION USING FIVE SELECTED VARIABLES\textsuperscript{a} TO PREDICT THERAPEUTIC OUTCOME
N=33

<table>
<thead>
<tr>
<th>Outcome Measures</th>
<th>r</th>
<th>r\textsuperscript{2}</th>
<th>F\textsubscript{5, 27}</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intelligence</td>
<td>.399</td>
<td>.159</td>
<td>1.02</td>
<td>N.S.</td>
</tr>
<tr>
<td>Classroom Achievement</td>
<td>.495</td>
<td>.245</td>
<td>1.75</td>
<td>N.S.</td>
</tr>
<tr>
<td>Reduction in Behavior Problems</td>
<td>.492</td>
<td>.242</td>
<td>1.72</td>
<td>N.S.</td>
</tr>
<tr>
<td>Achievement Motivation</td>
<td>.372</td>
<td>.139</td>
<td>.87</td>
<td>N.S.</td>
</tr>
<tr>
<td>Combined Index</td>
<td>.539</td>
<td>.291</td>
<td>2.21</td>
<td>N.S.</td>
</tr>
</tbody>
</table>

\textsuperscript{a}Criterion variables include late-rated nonpossessive warmth, late-rated accurate empathy, child relationship inventory total score, helper relationship inventory total score, and helper dogmatism score.

which can be accounted for. In general, it can be stated that the combined variables have been able to account for 29 per cent of the variance for the combined index, 24 per cent of the variance for change in achievement and reduction in behavior problems, 16 per cent of the variance in change in intelligence, and 14 per cent of the variance in change in motivation. Because of the small increase in correlation when predictive criteria are combined, it might be tentatively suggested that the predictive elements of one criterion tend to be common to another. Even in the case of the combined outcome index, 71 per cent of the variance has not
been accounted for. If the trend should continue in a larger sample, however, multiple correlations would become significant.

**Summary of Chapter IV**

Reliable ratings of interviews were obtained for nonpossessive warmth and accurate empathy. Ratings of these conditions in the present study tended to be lower than that reported by Rogers *et al.* (1967) in a similar study using trained therapists. Ratings of the late interview did not correlate highly with ratings of the early interview.

Helpers are more optimistic than children in their judgment of the relationship. Children's relationship inventories tend to reflect a global feeling about the relationship while helpers distinguish between variables. Dogmatism scores for helpers tend to be higher than those reported for other groups.

Rated conditions within the same interview tend to be related as are helper and child relationship inventories. However, rated conditions agree much less with helper and child perceptions. Dogmatism scores show no relation to measured conditions.

The data suggests that rated nonpossessive warmth, rated accurate empathy, and the child relationship inventory total score are related to case outcome. Predictive elements of one criterion appear to be common to another.
CHAPTER V

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

Chapter V presents the summary, conclusions and implications of the study. Suggestions for further study are also included.

Review of previous research suggests that there are elements within the therapeutic relationship which cut across theoretical boundaries. Three of these elements, (a) genuineness, (b) communication of nonpossessive warmth, and (c) communication of accurate empathy have been defined and investigated by client-centered therapists. The evidence suggests that therapeutic outcome is related to these three conditions for neurotic and schizophrenic adults. In addition, evidence is available which suggests that the open minded helping person is more likely to provide these three conditions and therefore be more therapeutic in the relationship.

In the present study, data were collected in an attempt to extend and clarify the evidence concerning the three therapeutic variables in a relationship involving untrained volunteers working with elementary school children who were experiencing behavioral and academic difficulty. In
addition, an attempt was made to show that open-minded volunteers would provide higher levels of therapeutic conditions than close-minded volunteers.

**Summary of Procedure**

The specific purposes of this study were to investigate the hypotheses that (1) high levels of genuineness, non-possessive warmth, and accurate empathy in the helping person as rated by judges are significantly related to therapeutic outcome, (2) high levels of these conditions as perceived by the helper and the child are significantly related to therapeutic outcome, and (3) open mindedness in the helping person is significantly related to therapeutic outcome.

The sample included 35 students and 35 adult helpers who participated in the Franklin County Community Helper Program during the 1966-67 school year. Community helpers were volunteers who worked with one child twice each week for approximately three months. Interviews varied from 30 minutes to one hour in length. The children who received the treatment had been diagnosed as academic and behavior problems in the school. The purpose of the Community Helper Program was to determine the value of using untrained personnel in the helping relationship with a minimum of professional supervision. Helpers were instructed to establish a "good" relationship. They were allowed to choose their activities and to structure the relationship in their own way. Each of the children who received treatment was examined before and
after the treatment program. Five indexes of success utilized in the Community Helper Program were made available for the present study. They included gains in intelligence, gains in achievement, reduction in classroom behavior problems, improved motivation as measured by picture projectives and total gain scores obtained by combining the four measures of success.

The instruments used in this study were: (1) A Scale for the Measurement of Unconditional Positive Regard, (2) A Scale for the Measurement of Accurate Empathy, (3) A Scale for the Measurement of Therapist Self-Congruence, (4) a relationship inventory adapted specifically for the study, and (5) the Rokeach Dogmatism Scale. An explanation of each instrument is included in Chapter III and sample copies are included in the appendices.

The first three instruments were used to rate the level of therapeutic conditions in helper-child interviews. Audio tapes of two helper-child interviews early in the treatment program and two interviews late in the program were available for analysis. From each of the second and fourth tapes, three three-minute segments were selected by a standardized procedure for analysis. Thus, 210 interview segments made up the data to be rated. Three trained judges listened to each segment and rated it according to the levels of the scale. Nonpossessive warmth, accurate empathy, and genuineness were rated in that order by the same three
judges. When reliability between judges was high for a particular scale, ratings were combined for further analysis. These combined levels of rated conditions were correlated with each of the five outcome indexes.

The relationship inventories were used to determine child and helper perceptions of the levels of conditions offered by the helper. Inventories were examined for intercorrelations among subscales. Total inventory scores were also correlated with levels of rated conditions and each of the five outcome indexes.

The Rokeach Dogmatism Scale was used to determine the relative dogmatism of each helper. Dogmatism scores were intercorrelated with levels of rated conditions, helper and child relationship inventory total scores, and each of the five outcome indexes.

In order to account for as much of the outcome variance as possible, multiple correlations were determined between five selected conditions measures and each of the five outcome indexes.

The Findings

Included in the findings of this study are the characteristics of measured conditions and relationship between measured levels of therapeutic conditions along with evidence directly related to each of the hypotheses.

The following conclusions seem warranted with respect to the characteristics of measured variables.
1. Without instruction, untrained volunteers who attempt to establish a "good" one-to-one relationship with a child tend to provide relatively low levels of genuineness, nonpossessive warmth, and accurate empathy as rated by judges.

2. Levels of nonpossessive warmth from ratings taken near the end of treatment cannot be predicted from ratings taken from the third and fourth interviews. Rated levels of accurate empathy also show considerable variation from the early to late interview but can be predicted more accurately than nonpossessive warmth.

3. As rated by judges, a helper who provides a high level of non-possessive warmth in an interview will also tend to provide a high level of accurate empathy within that same interview.

4. Helpers tend to perceive themselves in the helping relationship more favorably than do the children.

5. Children respond to the items of the relationship inventory in a global fashion rather than differentiating between variables. This global condition appears to be related to the helper's relationship inventory total score and to nonpossessive warmth as rated by judges.

6. Helper dogmatism scores are not related to any of the rated or perceived measures of levels of therapeutic conditions.
The following conclusions seem warranted with respect to Hypothesis 1:

a. Because judges were unable to agree upon the level of genuineness offered by a helper, it was impossible to relate this variable to therapeutic outcome.

b. Helpers who are judged to have provided high levels of nonpossessive warmth in a late interview bring about significant positive behavioral change as measured by gains in achievement, by reduction in behavior problems, and by gains reflected in the combined outcome index.

c. Helpers who are judged to have provided high levels of accurate empathy in a late interview bring about significant positive behavioral change as measured by gains in achievement and by gains reflected in the combined outcome index.

The following conclusions seem warranted with respect to Hypothesis 2:

a. The child's perception of the level of conditions offered by the helper and measured near the end of treatment is significantly related to reduction in behavior problems and gains reflected in the combined outcome index.

b. The helper's perception of the level of conditions offered by the helper and measured near the end of treatment is not significantly related to therapeutic outcome as reflected by the indexes used for this study.
The following conclusion seems warranted with respect to Hypothesis 3.

Helper dogmatism as measured by the Rokeach Scale is not related to therapeutic outcome as reflected by the indexes used for this study.

Multiple correlations utilizing the combined measures of rated conditions, perceived conditions, and helper dogmatism do not greatly increase the amount of outcome variance which can be accounted for.

Implications

The study has been concerned with attempting to extend and clarify the evidence which suggests that genuineness, nonpossessive warmth, and accurate empathy on the part of the helping person is related to positive behavioral change in the person being helped. Specifically, the relationship between an elementary school child and an adult volunteer is used to investigate the hypotheses concerning therapeutic outcome. From the evidence presented, the following implications seem warranted:

1. This study attempted to rate all helpers on three therapeutic variables. Rated nonpossessive warmth was so highly related to rated accurate empathy that little was gained by the second set of ratings. If future studies confirm these results, there would be little value in rating for both variables. Due to the inability of judges to agree,
this study was unable to investigate the importance of helper genuineness in the therapeutic interview.

2. The ratings suggest that helpers, in general, provide relatively low levels of therapeutic conditions when left to their own design. As a group, then, they do not tend to bring about significant positive behavioral change. Future programs of this nature should be concerned with finding ways of helping the volunteers to increase their levels of therapeutic conditions.

3. It has been shown that the child's perception of the relationship is significantly related to case outcome. These perceptions are easily obtained and should be used in future programs. The instrument used in this study needs considerable refinement. A reliable child relationship inventory might well prove the best predictor of therapeutic outcome.

4. The ratings and the child perceptions which had predictive value were acquired near the end of treatment. Their ultimate usefulness will be determined by the earliest date which they become consistent and related to outcome.

5. Helper Dogmatism Scales appear to be unrelated to therapeutic outcome or level of conditions offered by the helper. A method of predicting which volunteers will be most likely to provide high levels of therapeutic conditions is still needed.
Suggestions for Further Study

Studies are needed which will support, disprove, or further clarify the tentative conclusions and implications made from this study. The following suggestions are offered:

1. Further research should be conducted comparing levels of rated conditions with therapeutic outcome for other groups of children. A larger number of audio tapes taken throughout the treatment program would provide evidence as to the stability of the level of therapeutic conditions offered by the helper and the number of interviews needed before outcome prediction could take place.

2. If community helpers are to make a significant contribution to the total school program, certain steps must be taken to improve the average level of therapeutic conditions. One way this might be done is through better selection procedures. Future research should consider other instruments which may help to select helpers who will be able to provide high levels of therapeutic conditions. This study used all mothers. Future studies might consider using students, fathers, and retired persons. Finally, a more comprehensive school volunteer program utilizing helpers for various tasks might give professional persons a better chance to observe potential helpers before assigning them to the helping relationship.

There is evidence which suggests that lay personnel can learn to provide high levels of nonpossessive warmth and
accurate empathy in a relatively short period of time (Truax and Carkhuff, 1966). Future studies should investigate this and other ways of helping volunteers to raise the level of therapeutic conditions which they offer.

3. There is a need for a relationship inventory which will more effectively tap the child's perceptions of the conditions offered by the helper. It should be administered early in treatment and repeated several times throughout the program in order to discover the earliest point at which the inventory score becomes related to case outcome.

4. Future studies should not discard the possible value of helper perceptions of the level of conditions which they offer. Refinement of the present inventory might well lead to significant correlations between helper perceptions and case outcome.

5. This study has assumed that the helpers are primarily responsible for the level of conditions offered by the therapist. Yet, there is considerable evidence which suggests that the level of conditions are a result of dynamic interaction between helper and child (Rogers, et al., 1966). There is a need for a study to investigate characteristics within the child which are related to the level of conditions offered by the helper.

6. This study has assumed that all children respond equally to the conditions offered by the helper. Variables such as age of child, socio-economic level of child, and
type of behavior problem presented need to be investigated to determine what kinds of children respond best to each of the therapeutic variables.

7. The therapeutic variables used in this study account for a relatively small amount of the total variance. Future studies should continue to seek other elements which will add to the therapeutic value of the helping relationship.

The present chapter has included a summary of the study, together with a review of conclusions based on the data obtained in the study. Implications based on the conclusions were presented. Finally, suggestions for further study were postulated.
APPENDIX A
Number of the Interviews Used for Analysis, Number of the Tape Used for Analysis, and Total Number of Interviews for each case

<table>
<thead>
<tr>
<th>Case Number</th>
<th>Early Interviews</th>
<th>Late Interviews</th>
<th>Total Interviews</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Interview No.</td>
<td>Tape No.</td>
<td>Interview No.</td>
</tr>
<tr>
<td>1</td>
<td>4</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>1</td>
<td>25</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>4</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>8</td>
<td>4</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>1</td>
<td>18</td>
</tr>
<tr>
<td>10</td>
<td>4</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>11</td>
<td>4</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>12</td>
<td>4</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>13</td>
<td>4</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>14</td>
<td>4</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>15</td>
<td>3</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>16</td>
<td>4</td>
<td>2</td>
<td>16</td>
</tr>
<tr>
<td>17</td>
<td>4</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>18</td>
<td>4</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>19</td>
<td>4</td>
<td>2</td>
<td>24</td>
</tr>
<tr>
<td>20</td>
<td>4</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>21</td>
<td>5</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>22</td>
<td>4</td>
<td>2</td>
<td>18</td>
</tr>
<tr>
<td>23</td>
<td>4</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>24</td>
<td>4</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>25</td>
<td>5</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>26</td>
<td>4</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>27</td>
<td>4</td>
<td>2</td>
<td>19</td>
</tr>
<tr>
<td>28</td>
<td>4</td>
<td>2</td>
<td>15</td>
</tr>
<tr>
<td>29</td>
<td>4</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>30</td>
<td>3</td>
<td>1</td>
<td>24</td>
</tr>
<tr>
<td>31</td>
<td>4</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>32</td>
<td>4</td>
<td>2</td>
<td>25</td>
</tr>
<tr>
<td>33</td>
<td>4</td>
<td>2</td>
<td>14</td>
</tr>
<tr>
<td>34</td>
<td>4</td>
<td>2</td>
<td>21</td>
</tr>
<tr>
<td>35</td>
<td>4</td>
<td>2</td>
<td>24</td>
</tr>
</tbody>
</table>
A SCALE FOR THE MEASUREMENT OF
THERAPIST SELF-CONGRUENCE

Charles B. Truax and Walter A. Dickenson

STAGE 1

There is striking evidence of contradiction between the therapist's experiencing and his current verbalization. The content of what he says is obviously unrelated to what he is really feeling at that moment. The discrepancy may show in the way his voice quality does not match what he is saying.

STAGE 2

There is good evidence of contradiction between the therapist's experiencing and his current verbalization. The content of what he says does not appear related to what he is feeling at that moment, although the evidence is not as clear cut and obvious as in Stage 1. His voice quality may suggest that he is feeling something else besides what he is saying, but it might be difficult to decide just what.

STAGE 3

There is slight evidence of contradiction between the therapist's experiencing and his current verbalization. This can come from a variety of sources. It can come from a "professional" tone that has a rehearsed quality as if he always says that particular thing in that way because that is what a good doctor is supposed to feel and say. Or it can come
from a slight sense of strain as if the therapist does not really feel what he is trying to give the impression of feeling. It may seem as if the therapist is striving to be nicer than he really is, to really like the patient and express warmer feelings toward him than he really feels, or to suppress feelings which are arising in him to some degree which he does not want to feel or express. His interaction may have a slightly mechanical or wooden quality as if he were a little stiff or reserved. If any of the qualities of this stage are more than slightly noticeable, then the rating should be lower to the degree that they are more noticeable, then rating should be lower to the degree that they are more noticeably present.

STAGE 4

There are no negative cues suggesting a discrepancy between what he says and what he feels, but there are no positive cues indicative of really genuine response to the patient either. The responses are appropriate, they seem sincere, or at least do not seem insincere, and yet there does not seem to be any real involvement either. He listens and follows, and may even do his job well, but nothing more.

STAGE 5

There are no negative cues suggesting any discrepancy between what he says and what he feels, and there are some positive cues indicating genuine response to the patient. These genuine feelings can be positive or negative so long
as the words and content are not at odds with the feelings going with them. There is no mechanical response here; the therapist feels his words although he may be hesitant about expressing them fully. He may appear somewhat restrained or somewhat less enthusiastic, or frightened or angry, or depressed or whatever the patient is feeling, but his words sincerely express his own reaction. They do not seem at all phony, though they may be somewhat muted.

STAGE 6

The genuine response is clear and solid. The therapist responds with a goodly portion of his own feelings and there is no doubt as to whether he really means what he says. He may not be completely free in expressing them, he may have to pause and retreat on occasion but he does not hide his confusion or errors. He does not seem to be at a loss for words or to be in doubt about what he feels or should say. He is definitely in his words, and there is no mistaking his genuine reaction.

STAGE 7

The therapist is freely and deeply himself in the relationship. There is an openness to experiencing and feelings of all types, both positive and negative. He can laugh or cry, get thoroughly angry or sexually stimulated, be quietly accepting or very active. There may be contradictory feelings but these are accepted or recognized and not denied or fended off in any way. The therapist is clearly
being himself in all of his responses whether these responses are personally meaningful or trite, empathic or obtuse. His words clearly convey his most genuine feelings. The keynote is his spontaneous humanity. No example was included because these rare moments in therapy will be readily apparent when they occur.
A SCALE FOR THE MEASUREMENT OF
UNCONDITIONAL POSITIVE REGARD

Charles B. Truax

STAGE 1

The therapist is actively offering advice or giving clear negative regard. He may be telling the patient what would be "best" for him, or may be in other ways actively either approving or disapproving of his behavior. The therapist acts in such a way as to make himself the locus of evaluation. The therapist sees himself as responsible for the patient.

STAGE 2

The therapist responds mechanically to the client and thus indicates little positive regard and hence little unconditional positive regard. The therapist may ignore the patient or his feelings or display a lack of concern or interest for the patient. Therapist ignores client where an unconditional positive regard response would be expected--complete passivity that communicates almost unconditional lack of regard.

STAGE 3

The therapist indicates a positive caring for the patient or client but it is a semi-possessive caring in the sense that he communicates to the client that what the client does or does not do, matters to him. That is, he communicates
such things as "it is not all right if you act immorally," "I want you to get along at work," or "it's important to me that you get along with the ward staff." The therapist sees himself as responsible for the client.

STAGE 4

The therapist clearly communicates a very deep interest and concern for the welfare of the patient. The therapist communicates a nonevaluative and unconditional positive regard to the client in almost all areas of his functioning. Thus, although there remains some conditionality in the more personally and private areas the patient is given freedom to be himself and to be liked as himself. Thus, evaluating thoughts and behaviors are for the most part absent. In deeply personal areas, however, the therapist may be conditional so that he communicates to the client that the client may act in any way he wishes except that it is important to the therapist that he be more mature or that he not regress in therapy or that the therapist himself is accepted and liked. In all other areas, however, Unconditional Positive Regard is communicated. The therapist sees himself as responsible for the client.

STAGE 5

At Stage 5, the therapist communicates Unconditional Positive Regard without restriction. There is a deep respect for the patient's worth as a person and his rights as a free individual. At this level the patient is free to
be himself even if this means that he is regressing, being defensive, or even disliking or rejecting the therapist himself. At this stage, the therapist cares deeply for the patient as a person but it does not matter to him in which way the patient may himself choose to behave. There is a caring for and a prizing of the patient for his human potentials. This genuine and deep caring is uncontaminated by evaluations of his behavior or his thoughts. There is a willingness to equally share the patient's joys and aspirations or his depressions and failures. The only channeling by the therapist may be the demand that the patient communicate personally relevant material.
APPENDIX D
A SCALE FOR THE MEASUREMENT OF ACCURATE EMPATHY

Charles B. Truax

STAGE 1

Therapist seems completely unaware of even the most conspicuous of the client's feelings. His responses are not appropriate to the mood and content of the client's statements and there is no determinable quality of empathy, hence, no accuracy whatsoever. The therapist may be bored and disinterested or actively offering advice but he is not communicating an awareness of the client's current feelings.

STAGE 2

Therapist shows a degree of accuracy which is almost negligible in his responses, and then only toward the client's most obvious feelings. Any emotions which are not so clearly defined, he tends to ignore altogether. He may be correctly sensitive to obvious feelings and yet misunderstand much of what the client is really trying to say. By his response he may block off or may misdirect the patient. Stage 2 is distinguishable from Stage 3 in that the therapist ignores feelings rather than displaying an inability to understand feelings.

STAGE 3

Therapist often responds accurately to client's more exposed feelings. He also displays concern for the deeper, more hidden feelings, which he seems to sense must be present,
though he does not understand their nature. The therapist seems to assume the presence of deep feelings, although he does not sense their meaning to this particular patient.

**STAGE 4**

Therapist usually responds accurately to the client's more obvious feelings and occasionally recognizes some that are less apparent. In the process of this tentative probing, however, he may anticipate feelings. Sensitivity and awareness of the therapist are present but he is not entirely "with" the patient in the current situation or experience. (The desire and effort to understand are both present but accuracy is low.) It is distinguishable from Stage 2 in that the therapist does occasionally recognize feelings that are less apparent. Also the therapist may seem to have a theory about the patient and may even know how or why the patient feels a particular way, but the therapist is definitely not "with" the patient—they are not together. In short, the therapist may be diagnostically accurate, but not empathically accurate in his sensitivity to the current feeling stage of the patient.

**STAGE 5**

Therapist accurately responds to all of the client's more readily discernible feelings. He shows awareness of many feelings and experiences which are not so evident, too, but in these he tends to be somewhat inaccurate in his understanding. The therapist may recognize more feelings that
are not so evident. When he does not understand completely this lack of complete understanding is communicated without an anticipatory or jarring note. His misunderstandings are not disruptive by their tentative nature. Sometimes in Stage 5 the therapist simply communicates his awareness of the problem of understanding another person's inner world. Stage 5 is the midpoint of the continuum of accurate empathy.

**STAGE 6**

Therapist recognizes most of the client's present feelings, including those which are not readily apparent. Sometimes, however, he tends to misjudge the intensity of these veiled feelings, with the result that his responses are not always accurately suited to the exact mood of the client. In content, however, his understanding or recognition includes those not readily apparent. The therapist deals with feelings that are current with the patient. He deals directly with what the patient is currently experiencing although he may misjudge the intensity of less apparent feelings. Often the therapist, while sensing the feelings, is unable to communicate meaning to these feelings. The therapist statements contain an almost static quality in contrast to Stage 7 in the sense that the therapist handles those feelings that the patient offers but does not bring new elements to life. He is with the client but doesn't encourage
exploration. His manner of communicating his understanding is such that he makes of it a finished thing.

STAGE 7

Therapist responds accurately to most of the client's present feelings. He shows awareness of the precise intensity of most underlying emotions. However, his responses move only slightly beyond the area of the client's own awareness, so that feelings may be present which are not recognized by the client or therapist. The therapist may communicate simply that the patient and he are moving towards more emotionally significant material. Stage 7 is distinguishable from Stage 6 in that often the therapist response is a kind of pointing of the finger toward emotionally significant material with great precision in the direction of pointing.

STAGE 8

Therapist accurately interprets all the client's present, acknowledged feelings. He also uncovers the most deeply-shrouded of the client's feeling areas, voicing meanings in the client's experience of which the client is scarcely aware. Since he must necessarily utilize a method of trial and error in the new uncharted areas, there are resulting minor flaws in the accuracy of his understanding, but inaccuracies are held tentatively. He moves into feelings and experiences that are only hinted at by the client and does so with sensitivity and accuracy. The therapist offers
specific explanations or additions to the patient's understanding so that not only are underlying emotions pointed to, but they are specifically talked about. The content that comes to life may be new but is not alien. While the therapist in Stage 8 makes mistakes, mistakes do not have a jarring note, but are covered by the tentative character of the response. Also the therapist is sensitive to his mistakes and quickly alters or changes his response in midstream, indicating that he more clearly knows what is being talked about and what is being sought after in the patient's own explorations. The therapist reflects a togetherness with the patient in tentative trial and error exploration. His voice tone reflects the seriousness and depth of his empathic grasp.

**STAGE 9**

Therapist unerringly responds to the client's full range of feelings in their exact intensity. Without hesitation, he recognizes each emotional nuance and communicates an understanding of every deepest feeling. He is completely attuned to the client's shifting emotional content; he senses each of the client's feelings and reflects them in his words and voice. He expands the client's hint into a full-blown but tentative elaboration of feeling or experience with unerring sensitive accuracy. Both a precision in understanding and a precision in the communication of this understanding are present. Both are expressed and experienced by the therapist without hesitancy.
RELATIONSHIP INVENTORY—HELPER

People feel differently about some people than they do about others. There are a number of statements below that describe a variety of ways that one person may feel about another person, or ways that one person may act towards another person. Consider each statement carefully and decide whether it is true or false when applied to your present relationship with _________. If the statement seems mostly to be true then mark it true, while if it is mostly not true, then mark it false.

1. I like seeing the child.
2. Sometimes I like him better than at other times.
3. I understand his words but not the way he feels.
4. I pretend that I like him more than I really do.
5. I am friendly and warm toward him.
6. I like him in some ways. In other ways I do not.
7. I nearly always know exactly what he means.
8. Sometimes I say one thing and do something else.
10. When he gets mad at me, I get mad at him.
11. I understand him.
12. I get upset when he talks about certain things.
13. I just put up with him.
14. Sometimes I praise him, at other times I criticize him.
15. I know how he feels about things that happen to him.
16. I always try to be honest with him.
17. I don't really care what happens to him.
18. I like him better when he behaves.
19. I don't know what things are important to him.
20. I try not to say anything that might upset him.
22. I am upset when he doesn't do his very best.
23. I understand what he means even if he can't say it very well.
24. When I say something, I really mean it.
RELATIONSHIP INVENTORY--CHILD

People feel differently about some people than they do about others. Here are a number of statements that show ways that one person may feel about another person, or ways that one person may act toward another person. Listen to each statement carefully and decide whether it is true or false when applied to your present relationship with ________. If the statement seems to mostly be true then say "true." If it is mostly not true then say "false."

____ 1. She likes seeing me.
____ 2. Sometimes she likes me better than at other times.
____ 3. She understands my words but not the way I feel.
____ 4. She pretends that she likes me more than she really does.
____ 5. She is friendly and warm towards me.
____ 6. She likes me in some ways. In other ways she does not.
____ 7. She nearly always knows exactly what I mean.
____ 8. Sometimes she says one thing and does something else.
____ 9. She cares about me.
____ 10. When I get mad at her, she gets mad at me.
____ 11. She understands me.
____ 12. She gets upset when I talk about certain things.
____ 13. She just puts up with me.
____ 14. Sometimes she praises me, at other times she criticizes me.
____ 15. She knows how I feel about things that happen to me.
____ 16. She always tries to be honest with me.
____ 17. She doesn't really care what happens to me.
____ 18. She likes me better when I behave.
____ 19. She doesn't know what things are important to me.
____ 20. She tries not to say anything that might upset me.
____ 21. She dislikes me.
____ 22. She is upset when I don't do my very best.
____ 23. She understands what I mean even if I can't say it very well.
____ 24. When she says something, she really means it.
ROKEACH DOGMATISM SCALE

The following is a survey of the opinions of people in general about a number of social and personal questions. Of course there are many different answers. The best answer to each statement below is your personal opinion. We have tried to cover many different and opposing points of view; you may find yourself agreeing strongly with some of the statements disagreeing just as strongly with others, and perhaps uncertain about others; whether you agree or disagree with any statement, you can be sure that many other people feel the same as you do.

Mark each statement in the left margin according to how much you agree or disagree with it. Please mark every one. Write +1, +2, +3, or -1, -2, or -3 depending on how you feel in each case.

+1: I agree a little  
+2: I agree on the whole  
+3: I agree very much  
-1: I disagree a little  
-2: I disagree on the whole  
-3: I disagree very much

1. A person who thinks primarily of his own happiness is beneath contempt.
2. The main thing in life is for a person to want to do something important.
3. I wish people would be more definite about things.
4. In a discussion I often find it necessary to repeat myself several times to make sure I am be understood.
5. Most people just don't know what's good for them.
6. A person who has bad manners, habits, and breeding can hardly expect to get along with decent people.
7. In times like these, a person must be pretty selfish if he considers his own happiness primarily.
8. A man who does not believe in some great cause has not really lived.
9. I work under a great deal of tension at times.
10. I'd like it if I should find someone who would tell me how to solve my personal problems.
11. Of all the different philosophies which have existed in this world there is probably only one which is correct.

12. Whether it's all right to manipulate people or not, it is certainly all right when it's for their own good.

13. It is when a person devotes himself to an ideal or cause that his life becomes meaningful.

14. In this complicated world of ours the only way we can know what is going on is to rely upon leaders or experts who can be trusted.

15. If people would talk less and work more, everybody would be better off.

16. There are a number of persons I have come to hate because of the things they stand for.

17. There is so much to be done and so little time to do it in.

18. It is when a person devotes himself to an ideal or cause that he becomes important.

19. It is better to be a dead hero than a live coward.

20. A group which tolerates too much difference of opinion among its own members cannot exist for long.

21. The businessman and the manufacturer are much more important to society than the artist and the professor.

22. It is only natural that a person should have a much better acquaintance with ideas he believes in than with ideas he opposes.

23. While I don't like to admit this even to myself, I sometimes have the ambition to become a great man like Einstein, or Beethoven, or Shakespeare.

24. Plain common sense tells you that prejudice can be removed by education, not legislation.

25. Even though freedom of speech for all groups is worthwhile goal, it is unfortunately necessary at times to restrict the freedom of certain political groups.
26. If a man is to accomplish his mission in life it is sometimes necessary to gamble "all or nothing at all."

27. A person must be pretty stupid if he still believes in differences between races.

28. Most people just don't give a "damn" about others.

29. A person who gets enthusiastic about a number of causes is likely to be a pretty "wisy'washy" sort of person.

30. Do unto others ad they do unto you.

31. To compromise with our political opponents is dangerous because it usually leads to the betrayal of our own side.

32. If given the chance I would do something that would be of great benefit to the world.

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

34. In times like these it is often necessary to be more on guard against ideas put out by certain people or groups in one's own camp than by those in the opposing camp.

35. In a heated discussion I generally become so absorbed in what I am going to say that I forget to listen to what the others are saying.

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

37. Once I get wound-up in a heated discussion I just can't stop.

38. There are two kinds of people in this world; those who are on the side of truth and those who are against it.

39. What the youth needs is strict discipline, rugged determination, and the will to work and fight for family and country.

40. Man on his own is a helpless and miserable creature.

41. The United States and Russia have just about nothing in common.
42. I set a high standard for myself and I feel others should do the same.

43. In the history of mankind there have probably been just a handful of really great thinkers.

44. The highest form of government is a democracy and the highest form of democracy is a government run by those who are most intelligent.

45. Appreciation of others is a healthy attitude, since it is the only way to have them appreciate you.

46. The present is all too often full of unhappiness. It is the future that counts.

47. Unfortunately, a good many people with whom I have discussed important social and moral problems don't really understand what is going on.

48. People who seem unsure and uncertain about things make me feel uncomfortable.

49. Fundamentally, the world we live in is a pretty lonely place.

50. It is often desirable to reserve judgment about what's going on until one has had a chance to hear the opinions of those one respects.

51. In general, full economic security is bad; most men wouldn't work if they didn't need the money for eating and living.

52. The worst crime a person can commit is to attack publicly the people who believe in the same thing he does.

53. In the long run the best way to live is to pick friends and associates whose tastes and beliefs are the same as one's own.

54. The American re-armament program is clear and positive proof that we are willing to sacrifice to preserve our freedom.

55. Most of the ideas which get published nowadays aren't worth the paper they are printed on.

56. It is only natural for a person to be rather fearful of the future.
57. Most of the arguments or quarrels I get into are over matters of principal.

58. My blood boils whenever a person stubbornly refuses to admit he's wrong.

59. When it comes to differences of opinion in religion we must be careful not to compromise with those who believe differently from the way we do.

60. America may not be perfect, but the American way has brought us about as close as human beings can get to a perfect society.

Name ____________________________
### MEAN VALUES AND STANDARD DEVIATIONS OF THE CHILD AND HELPER RELATIONSHIP INVENTORY SUBSCALES

**N=35**

<table>
<thead>
<tr>
<th>Subscale</th>
<th>Child RI Mean Scores</th>
<th>Child RI Standard Deviations</th>
<th>Helper RI Mean Scores</th>
<th>Helper RI Standard Deviations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Regard</td>
<td>5.31</td>
<td>.91</td>
<td>5.91</td>
<td>.27</td>
</tr>
<tr>
<td>Accurate Empathy</td>
<td>4.65</td>
<td>1.56</td>
<td>3.91</td>
<td>1.85</td>
</tr>
<tr>
<td>Genuineness</td>
<td>3.82</td>
<td>1.27</td>
<td>4.79</td>
<td>.52</td>
</tr>
<tr>
<td>Conditionality of Regard</td>
<td>2.11</td>
<td>1.48</td>
<td>3.85</td>
<td>1.45</td>
</tr>
<tr>
<td>Total Score</td>
<td>15.91</td>
<td>4.13</td>
<td>18.48</td>
<td>2.80</td>
</tr>
</tbody>
</table>

### PEARSON PRODUCT-MOMENT INTERCORRELATIONS AMONG THE CHILD RELATIONSHIP INVENTORY SUBSCORES

**N=35**

<table>
<thead>
<tr>
<th>Subscales</th>
<th>Accurate Empathy</th>
<th>Genuineness</th>
<th>Conditionality of Regard</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Regard</td>
<td>.511</td>
<td>.289</td>
<td>.328</td>
<td>.623</td>
</tr>
<tr>
<td>Accurate Empathy</td>
<td></td>
<td></td>
<td>.519</td>
<td>.837</td>
</tr>
<tr>
<td>Genuineness</td>
<td></td>
<td></td>
<td>.642</td>
<td>.798</td>
</tr>
<tr>
<td>Conditionality of Regard</td>
<td></td>
<td></td>
<td></td>
<td>.827</td>
</tr>
</tbody>
</table>
PEARSON PRODUCT-MOMENT INTERCORRELATIONS AMONG THE HELPER RELATIONSHIP INVENTORY SUBSCORES
N=35

<table>
<thead>
<tr>
<th>Subscales</th>
<th>Accurate Empathy</th>
<th>Genuineness</th>
<th>Conditionality of Regard</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Regard</td>
<td>.205</td>
<td>-.311</td>
<td>.180</td>
<td>.271</td>
</tr>
<tr>
<td>Accurate Empathy</td>
<td>-.017</td>
<td>.301</td>
<td></td>
<td>.837</td>
</tr>
<tr>
<td>Genuineness</td>
<td></td>
<td>.037</td>
<td></td>
<td>.163</td>
</tr>
<tr>
<td>Conditionality of Regard</td>
<td></td>
<td></td>
<td></td>
<td>.744</td>
</tr>
</tbody>
</table>

PEARSON PRODUCT-MOMENT INTERCORRELATIONS OF CHILD AND HELPER RELATIONSHIP INVENTORY SCORES
N=35

<table>
<thead>
<tr>
<th>(child) Subscale</th>
<th>(child) Level of Regard</th>
<th>Accurate Empathy</th>
<th>Genuineness</th>
<th>Conditionality of Regard</th>
<th>Total Score</th>
<th>(helper) Level of Regard</th>
<th>Accurate Empathy</th>
<th>Genuineness</th>
<th>Conditionality of Regard</th>
<th>Total Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level of Regard</td>
<td>-.228</td>
<td>.233</td>
<td>.011</td>
<td>.289</td>
<td>.284</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Accurate Empathy</td>
<td>.063</td>
<td>.372</td>
<td>.055</td>
<td>.466</td>
<td>.506</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Genuineness</td>
<td>.118</td>
<td>.295</td>
<td>.076</td>
<td>.309</td>
<td>.382</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Conditionality of Regard</td>
<td>.160</td>
<td>.520</td>
<td>.175</td>
<td>.534</td>
<td>.671</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total Score</td>
<td>.067</td>
<td>.471</td>
<td>.110</td>
<td>.529</td>
<td>.615</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
BIBLIOGRAPHY


Hirshberg, F., Carkhuff, R. R., and Berenson, B. G. "The Differential Effectiveness of Counselors and Therapists with In-patient Schizophrenics and Counseling Center Clients." Discussion Papers, Arkansas Rehabilitation Research and Training Center, University of Arkansas, 1966, 1 (Whole No. 6).


Van Der Veen, F. "Perceived Therapist Conditions and Degree of Disturbance." Unpublished research report, University of Wisconsin Psychiatric Institute, 1962.