INFORMATION TO USERS

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

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

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

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

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

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

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

Xerox University Microfilms
300 North Zeeb Road
Ann Arbor, Michigan 48106
75-19,505

YOUNG, Robert Thomas, 1945-
  PREDICTING THE EFFECTIVENESS OF THERAPEUTIC
  TUTORS WITH PRIMARY SCHOOL DISADVANTAGED CHILDREN.

The Ohio State University, Ph.D., 1975
Psychology, clinical

Xerox University Microfilms, Ann Arbor, Michigan 48106

THIS DISSERTATION HAS BEEN MICROFILMED EXACTLY AS RECEIVED.
PREDICTING THE EFFECTIVENESS OF THERAPEUTIC TUTORS WITH PRIMARY SCHOOL DISADVANTAGED CHILDREN

DISSERTATION

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

By

Robert T. Young, B.A., M.A.

* * * * * *

The Ohio State University
1975

Approved by

[Signature]
Adviser
Department of Psychology
ACKNOWLEDGMENTS

Considerable thanks must go to Dr. Herbert Rie of the Children's Hospital and The Ohio State University, who developed and directed the therapeutic tutoring program with which this study was concerned. He provided valued support and helpful criticisms to the author throughout his graduate studies, helped to shape some of his thinking, and served on the reading committee of this dissertation.

I would also like to thank Dr. Malcolm Helper, who served as my principal adviser throughout most of my graduate training, for his interest in my professional development and for his helpful suggestions concerning this study. Dr. Lyle Schmidt deserves special thanks for his constructive remarks concerning this dissertation and for agreeing so readily to serve on the reading committee of a student with whom he had little previous contact.

Thanks is also extended to the Columbus Public School System and specifically Livingston Avenue, Moler, Fair, Fairwood, Heyl, Reeb, Main, and Hubbard Elementary schools for supporting this program by providing the children and the required space. Dr. Ellen Rie is thanked for allowing the author to use her Parent-Teacher Questionnaire and Susie Newkirk and Sol Levin are thanked for conducting the parent interviews.

My greatest appreciation however is extended to my wife, Teena, for the love and support she gave so willingly throughout the period of my graduate studies. At times when things seemed most difficult her encouragement provided the needed inspiration to continue. Her interest
in this study, her help in rating the tapes, and her typing of the manuscript is also greatly appreciated.
PREFACE

This investigation was done in conjunction with an ongoing therapeutic tutoring service program which is one of several projects undertaken by the section on Child Development and Psychology at the Children's Hospital, Columbus, Ohio. These projects are funded by federal grant #NIH 76A105. This grant supports several projects in several different agencies and institutions and is entitled "Program for Linkage, Early Intervention and Prevention." Dr. Herbert E. Rie, Chief Psychologist, Children's Hospital, is the director of the programs that are based at the hospital. These projects began in 1972 and will run through 1980.
VITA

August 6, 1945 . . . . . . . . Born — Tucson, Arizona

1968 . . . . . . . . . . . . . . B.A., Hanover College, Hanover, Indiana

1968-1970 . . . . . . . . . Teaching Associate and Research Assistant, Department of Psychology, The Ohio State University, Columbus, Ohio

1970-1971 . . . . . . . . . Behavior Modification Specialist and Director of Education, Community Behavior Modification Programs, Columbus, Ohio

1971 . . . . . . . . . . . . . . M.A., The Ohio State University, Columbus, Ohio


1972-1975 . . . . . . . . . Research Associate, Children's Hospital, Columbus, Ohio

FIELDS OF STUDY

Major Field: Clinical Child Psychology

Minor Field: Developmental Psychology
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
</tr>
<tr>
<td>PREFACE</td>
</tr>
<tr>
<td>VITA</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
</tr>
<tr>
<td>INTRODUCTION</td>
</tr>
</tbody>
</table>

**Chapter**

<table>
<thead>
<tr>
<th></th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. REVIEW OF THE LITERATURE</td>
<td>5</td>
</tr>
<tr>
<td>II. METHOD</td>
<td>83</td>
</tr>
<tr>
<td>III. RESULTS</td>
<td>116</td>
</tr>
<tr>
<td>IV. DISCUSSION</td>
<td>155</td>
</tr>
</tbody>
</table>

**APPENDICES**

<table>
<thead>
<tr>
<th></th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>172</td>
</tr>
<tr>
<td>B</td>
<td>177</td>
</tr>
<tr>
<td>C</td>
<td>180</td>
</tr>
<tr>
<td>D</td>
<td>182</td>
</tr>
<tr>
<td>E</td>
<td>184</td>
</tr>
<tr>
<td>F</td>
<td>185</td>
</tr>
<tr>
<td>G</td>
<td>189</td>
</tr>
<tr>
<td>H</td>
<td>192</td>
</tr>
<tr>
<td>I</td>
<td>196</td>
</tr>
<tr>
<td>J</td>
<td>197</td>
</tr>
<tr>
<td>---</td>
<td>-----</td>
</tr>
<tr>
<td>K</td>
<td>200</td>
</tr>
<tr>
<td><strong>BIBLIOGRAPHY</strong></td>
<td>201</td>
</tr>
<tr>
<td>Table</td>
<td>Description</td>
</tr>
<tr>
<td>-------</td>
<td>-----------------------------------------------------------------------------</td>
</tr>
<tr>
<td>1.</td>
<td>CHILDREN'S GRADE LEVEL CATEGORIZED BY TUTORS (PA AND TCA GROUPS)</td>
</tr>
<tr>
<td>2.</td>
<td>AGE, RACE, EDUCATIONAL LEVEL, MEAN FAMILY INCOME AND INTELLIGENCE SCORES FOR THE TUTORS</td>
</tr>
<tr>
<td>3.</td>
<td>CHILDREN'S DEMOGRAPHIC AND INTELLIGENCE DATA (PA GROUP) CATEGORIZED BY TUTORS</td>
</tr>
<tr>
<td>4.</td>
<td>THE MEAN NUMBER OF SESSIONS AND THE MEAN NUMBER OF HOURS OF TUTORING RECEIVED PER CHILD</td>
</tr>
<tr>
<td>5.</td>
<td>MEAN TUTOR RATINGS OF THE REPRESENTATIVENESS OF AUDIO-TAPED SESSIONS</td>
</tr>
<tr>
<td>6.</td>
<td>INTERCORRELATIONS BETWEEN THE SEGMENTS FOR THE THERAPEUTIC CONDITIONS CATEGORIZED BY EACH RATER</td>
</tr>
<tr>
<td>7.</td>
<td>INTER-RATER RELIABILITY SCORES AND MEAN SUMMED CONDITION SCORES CATEGORIZED BY RATERS</td>
</tr>
<tr>
<td>8.</td>
<td>INTER-CORRELATION OF COMBINED RATER SUMMED CONDITION SCORES</td>
</tr>
<tr>
<td>9.</td>
<td>PRE-TUTORING MEANS AND MEAN CHANGE SCORES FOR THE OUTCOME VARIABLES</td>
</tr>
<tr>
<td>10.</td>
<td>MEAN OUTCOME CHANGE SCORES FOR EACH TUTOR</td>
</tr>
<tr>
<td>12.</td>
<td>CORRELATIONS BETWEEN THE MEAN NUMBER OF TUTORING SESSIONS RECEIVED PER CHILD AND THE TOTAL MEAN HOURS TUTORED PER CHILD WITH THE OUTCOME CHANGE SCORES</td>
</tr>
<tr>
<td>13.</td>
<td>INTERCORRELATIONS OF THE OUTCOME CHANGE SCORES</td>
</tr>
<tr>
<td>14.</td>
<td>ANALYSES OF COVARIANCE—TUTORS WITH THE OUTCOME CHANGE SCORES (WITH THEIR PRE SCORES COVARED)</td>
</tr>
</tbody>
</table>
## LIST OF TABLES (CONT.)

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.</td>
<td>TUTOR'S MEAN COMBINED RATER SUMMED CONDITION SCORES</td>
<td>127</td>
</tr>
<tr>
<td>16.</td>
<td>CORRELATION OF COMBINED RATER SUMMED CONDITION SCORES WITH THE OUTCOME CHANGE SCORES</td>
<td>133</td>
</tr>
<tr>
<td>17.</td>
<td>TUTOR PSYCHOMETRIC DATA INCLUDING THEIR INTELLIGENCE SCORES</td>
<td>135</td>
</tr>
<tr>
<td>18.</td>
<td>CORRELATIONS BETWEEN THE TUTOR PSYCHOMETRIC DATA AND THE MEAN TUTOR OUTCOME CHANGE SCORES</td>
<td>140</td>
</tr>
<tr>
<td>19.</td>
<td>UNIVARIATE ANALYSIS OF VARIANCE BETWEEN THE TUTORS AND THE THERAPEUTIC CONDITION RATINGS</td>
<td>143</td>
</tr>
<tr>
<td>20.</td>
<td>INTERCORRELATION TABLE OF THE COMBINED RATER SUMMED CONDITION SCORES AND THE TUTOR PSYCHOMETRIC SCORES</td>
<td>145</td>
</tr>
<tr>
<td>21.</td>
<td>THE RELATIONSHIP BETWEEN TUTOR DEMOGRAPHIC VARIABLES AND THE OUTCOME CHANGE SCORES</td>
<td>147</td>
</tr>
<tr>
<td>22.</td>
<td>THE RELATIONSHIP BETWEEN CHILD CHARACTERISTICS AND THE OUTCOME CHANGE SCORES</td>
<td>149</td>
</tr>
<tr>
<td>Figure</td>
<td>Description</td>
<td>Page</td>
</tr>
<tr>
<td>--------</td>
<td>-------------------------------------------------------------------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>1.</td>
<td>PERCENTAGES OF THE RATINGS ASSIGNED TO EACH STAGE OF THE ACCURATE EMPATHY SCALE FOR THE THERAPEUTIC TUTORING PROGRAM, THE WISCONSIN STUDY AND STOPPER'S COMMUNITY HELPER PROGRAM</td>
<td>128</td>
</tr>
<tr>
<td>2.</td>
<td>PERCENTAGES OF THE RATINGS ASSIGNED TO EACH STAGE OF THE NONPOSSESSIVE WARMTH SCALE FOR THE THERAPEUTIC TUTORING PROGRAM, THE WISCONSIN STUDY AND STOPPER'S COMMUNITY HELPER PROGRAM</td>
<td>130</td>
</tr>
<tr>
<td>3.</td>
<td>PERCENTAGES OF THE RATINGS ASSIGNED TO EACH STAGE OF THE GENUINENESS SCALE FOR THE THERAPEUTIC TUTORING PROGRAM AND THE WISCONSIN STUDY</td>
<td>131</td>
</tr>
</tbody>
</table>
INTRODUCTION

Man has long sought to achieve his fullest potential and to find his own sense of personal fulfillment. However, he has realized that few achieve this goal and consequently has attempted to determine, through scientific study, the variables that shape and mold his development. It is his hope, that by attaining a better understanding of the complex developmental process, he will achieve the knowledge that will enable mankind to come closer to realizing his fullest potential, intellectually and emotionally.

In the spirit of this quest for fulfillment and improvement, many have assumed academic enrichment to be a desirable goal and have proceeded to develop and implement programs aimed at facilitating children's academic performance. A large volume of literature has been published containing both theoretical explanations of the causes, and remedial techniques for overcoming the effects of learning disabilities, underachievement, and a disadvantaged social status. This study will concern itself with another approach to the problem of the poor academic achievement of disadvantaged children: therapeutic tutoring.

Therapeutic tutoring attempts to help a child improve his school achievement by concentrating not only on his specific subject matter, but also on the feelings and attitudes that are associated with his poor performance in the classroom. It is assumed that for many poorly achieving children their anxious expectations and confused impressions
of their school and their teachers are highly related to their poor school performance and that these feelings may actively interfere with their ability to learn. The longer these negative conditions exist, the further the child lags academically and additional defensive behaviors and attitudes accrue which lead to a further deterioration of their academic performance. Therapeutic tutoring recognizes the reciprocal interaction between the cognitive and emotional nature of the child and their combined effect on his academic performance. The therapeutic tutor attempts to remediate this poor performance by helping the child confront both his interfering feelings, attitudes, and expectations, and his specific subject matter simultaneously. Through the strength of the tutor-child relationship and the effective use of basic therapeutic principles the tutor helps the child disconfirm his learning inhibition and thereby enables him to make more effective use of his academic abilities. In essence, therapeutic tutoring combines aspects of both the therapeutic and educational functions, but retains as its primary emphasis educational enrichment.

This study is an attempt to determine the characteristics of successful tutors and of successful tutor-child relationships. It strives to relate some of the ingredients of the therapeutic tutoring relationship to the outcome variables, rather than focusing solely on the question of whether therapeutic tutoring is an effective process. Its focus is on discovering and isolating those elements within the therapeutic tutoring relationship that lead to constructive changes in the child's academic and behavioral performance. That is, it attempts to discover which elements, if any, among all those that are present,
contribute to constructive changes in the child's academic and behavioral performance.

If therapeutic tutoring and indeed all counseling and psychotherapy is to be an effective tool in the remediation of human malfunctioning, researchers and practitioners alike must attempt to discern the elements of any therapeutic relationship that inhibit constructive changes. It is now becoming clear from the research evidence that not everything that happens in a therapeutic relationship is beneficial or even relevant to constructive personality change. In fact, some studies indicate that in some cases the therapeutic relationship may prove detrimental to the client.

Some research evidence suggests that therapy is not really superior to receiving no treatment whatever. Eysenck (1961) reviewed nineteen reports covering more than 7,000 cases and concluded 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). Thus, many studies involving large numbers of subjects tend to suggest that the overall average effect of therapeutic interventions are approximately equivalent to the randomized effects of normal day to day living.

However, despite the evidence that average therapy is no better than no treatment, some studies involving specific therapists have produced positive and desireable changes in their subjects. It appears that some therapists are effective and helpful while others are ineffective and even harmful. The overall average effect, then, when all therapists
are considered, is comparable to receiving no help whatever.

If this conclusion is valid, the tremendous importance of determining for training, evaluation, professional, and licensing purposes the characteristics of the effective therapeutic relationship become obvious. Indeed, one could argue that all psychotherapy research should be concerned only with identifying those elements in the therapeutic process that relate to client outcomes. This study accepts that premise and attempts to specify the characteristics or ingredients in the therapeutic tutoring process that result in positive, neutral or deteriorative changes in the academic and behavioral performance of first, second, and third grade inner-city poor and underachieving disadvantaged children.
CHAPTER I

REVIEW OF THE RELATED LITERATURE

This review of the literature will be presented in nine sections. They include: 1) the underachieving child; 2) achievement and the disadvantaged child; 3) remedial programs for the disadvantaged, underachieving child; 4) the evidence of the overall effectiveness of counseling and psychotherapy; 5) the primary ingredients of effective counseling and psychotherapy: some theoretical views; 6) the meaning of accurate empathy, genuineness, and warmth and the development of scales to measure these therapeutic conditions; 7) the primary ingredients of effective counseling and psychotherapy: the empirical data; 8) the assessment of accurate empathy, non-possessive warmth and genuineness: methodological considerations; 9) predicting clinical effectiveness: a review of the literature on the prediction of outcome effectiveness from non-process variables.

THE UNDERACHIEVING CHILD

During the last decade vast quantities of research have explored the problem of academic underachievement. Researchers interested in this topic have attempted to ascertain the characteristics that would differentiate the underachiever from the student who is functioning at a level commensurate with his intellectual abilities (Teigland, et al., 1966). Traditionally, underachievement has implied a discrepancy between one's ability and one's achievement. This study however will be concerned primarily with culturally related underachievement in disadvantaged children and will not consider organic or physiological etiologies.

Although most of the underachievement literature is concerned with
college or secondary school students, Barrett (1957) and Shaw and McCuen (1960) have demonstrated the existence of the problem in elementary school children. In an Ohio study, DiSmye (1963) found that 16% of a representative cross section of the state's gifted elementary school children were not doing as well as expected. A conference on secondary schools (NEA, 1958) reported that 15 to 25% of the gifted students in most school systems are underachievers.

The extensive body of literature describing the characteristic behavior and development of underachievers contains evidence that underachievement is clearly related to a multitude of factors which affect and mold the child's development, i.e., academic, parent-child relationships, social acceptance, self-concept, etc. (Jackson, 1966).

Passow and Goldberg (1958) found that personal-social problems were related to underachievement. Vacc (1968), Granzow (1954) and Kurtz and Swenson (1951) found that underachievers were not as highly regarded as other children. The latter researchers also suggested that the peer group itself may facilitate and encourage poor academic performance. Further, Kurtz and Swenson found that underachievers did have friends, but they tended to have an unfavorable attitude toward school and a poor scholastic record. Teigland, et al. (1966) found that regardless of sex, underachievers are selected less frequently by their peers.

Rushton (1966), Blackham (1955), Granzow (1954), Vacc (1968), Norman and Daley (1959), Stone and Rowley (1964), Zimmerman (1965), and Yellott, et al. (1967) and others have found poor and underachievers to be poorly adjusted, immature, emotionally unstable, nervous and usually to possess feelings of hopelessness, discouragement and ambivalence. Bledsoe (1964), Buick and Bodurn (1962), Lumpkin (1959), and Shaw, Edson
and Bell (1960) all found underachievers to have poor self concepts, especially their male subjects.

Sears (1940), Calhoun (1956), and Sontag and Kagan (1963) all argue that underachievement is related to one's level of aspiration and one's desire for task mastery and intellectual competence. They suggest that those children who continue to fail, eventually set their aspirations with little regard for achievement and inevitably ultimately withdraw from all academic endeavors.

Abrams (1956) found that disabled readers, when compared to achieving adequate readers, were anxious, had difficulty in sustaining attention to abstract stimuli and were insecure and irritable. These findings lead them to conclude that these children were aware of their disability. In addition, these non-readers were found to be more impulsive and less able to respond appropriately to emotional stimuli. McMurray's (1963) results essentially agreed with those of Abrams. He found the poor reader to lack energy, be distractible, more likely to daydream, unable to complete assignments, very seldom relaxed, and very often to have repeated at least one grade in school.

Many investigators have attempted to relate varied parental characteristics to the underachievement of their children. Kurtz and Swenson (1951), Jackson (1968), Peppin (1963), Shaw and Dutton (1962), Morrow and Wilson (1961), Hilliard and Roth (1969) and others have indeed found that the nature of the parent-child relationship is related to a child's ability to achieve. Parents of underachievers were found to express more negative attitudes toward their children and usually tended not to expect much from them. On the other hand, the parents of achievers seem more approving and trusting in their relationship with their children,
more affectionate, and usually more encouraging of goal-oriented behaviors. In addition, several studies found that achievers' parents were more likely to read, play, and work with their children and also were more likely to share ideas and confidences with them.

A number of researchers, including Baer (1958), Carter (1956-57), and Norman, Clark and Bessener (1962) have concluded that overplacement is a significant factor in the development of the underachievement syndrome. Underachievers are often found to be among the youngest children in a certain grade and often a greater incidence of males is noted.

Shaw and McCuen (1960), Frankel (1960), Barrett (1957), and Farguhar and Taylor (1966), all concluded that underachievement begins rather early in a child's scholastic career and that the underachiever becomes progressively further and further behind as he continues to progress through his formal academic training.

After reviewing the literature concerned with the problem of underachievement one can draw a number of conclusions. Certainly this problem is complex and widespread, affecting millions of our children, who are not performing at a level commensurate with their abilities. This lost potential has obvious implications on our national collective potential and on our future growth and development. Unfortunately, the causes of poor and underachievement seem to be related to a multitude of factors and may vary from child to child, making remediation more difficult and costly. A few of these factors seem worthy of mention, however, by way of summarizing some of our current knowledge of this problem. Researchers have discovered that the underachieving child often is poorly
adjusted, has personal-social problems with his peers and parents, is often emotionally immature, has a poor self concept, is often highly distractable, and usually has little or no achievement motivation. With the presence of such severe psychological problems so clearly demonstrated, one can easily see how traditional therapeutic principles might be employed to aid in the remediation of the underachievement syndrome.

One is also struck by the many similarities between this literature and the literature on the poor academic achievement of the disadvantaged child.

ACHIEVEMENT AND THE DISADVANTAGED CHILD

During the last several decades a large number of investigators have become interested in studying the intellectual development of disadvantaged children in an attempt to determine the effect that cultural and sensory deprivation has on their cognitive development.

Many early investigations that were interested in exploring the effects of cultural or sensory deprivation focused on determining the existence of a critical period, during which the future development of children could be influenced with greater ease and with greater consequences than at any subsequent period (Lorenz, 1937; Scott and Marston, 1950; Harlow, 1949; Riesen, 1958; Levine, 1966). The critical period research has helped extend our knowledge of the results of sensory deprivation and the effect that selective stimulation plays in the cognitive and emotional development of a variety of animals. This data does seem to suggest that the earlier the intervention begins, the better the results.

"In the deprived home there seems to be a paucity of concern for the infant with steadily decreasing interest in the child, especially by the
time he reaches the toddler stage " (Deutsch, 1965, p. 80). Hunt (1966) observed that during the second year, stimulation of a verbal nature is lacking in culturally deprived homes.

Moss (1967) found that lower class parents engage in less face-to-face vocalization with their child during non-feeding times than do middle class mothers, who will hover over their children and play vocal games much more frequently. Bloom, Davis, and Hess (1965) concluded that the poor educational attainments of the parents, the frequent absence of the male parent, and the lack of parent child interaction and organized family activities all conspire to reduce the stimulation, language development, and intellectual development in disadvantaged homes.

Granzow's (1954) results support the findings of Bloom, et.al. (1965). He found that the majority of his underachievers in reading come from homes of lower socio-economic status and that their parents were less educated than the parents of achieving readers. In addition, these children often came from highly disrupted families where only one parent figure was present.

Curry (1962) found that as intellectual ability decreases, the consequent effect of social and economic conditions on scholastic achievement increases greatly. Those of low intellectual ability and low social class were found to have the poorest achievement levels, especially in reading. This same group was also found to have the poorest language skills when comparisons were made with more advantaged peers. Their data led the authors to suggest that the slower child may be the most vulnerable to the effects of cultural deprivation.
Riessman (1962) and Mingione (1965) found poor motivation for intellectual or school related activities in disadvantaged children. The cognitive style of low-income families seems qualitatively different from that found in middle class families and less effective in preparing the child for school and academic experiences.

Kirk (1966), Deutsch (1965), Mingione (1965), Klineberg (1963), Osborne (1960), and several others have suggested that the disadvantaged child tends to show a drop in rate of cognitive development in that they become progressively less able to handle intellectual and linguistic tasks (cumulative deficit phenomenon). Loretan (1966) argues that losses during the early formative years, due to an improper environment, may have an irreversible affect on cognitive development.

It appears that deprived or disadvantaged families differ from middle and upper class families in their educational aspirations, their verbal interaction with their child, the quality and quantity of stimulation available, their ability to provide the basic necessities of life, and in their level of academic achievement.

Currently, many of the differences found between lower and middle class children in their levels of achievement and their cognitive development are thought to result from cultural, maternal, and sensory deprivation. The problem of the disadvantaged child is now being viewed in such a way so as to attempt to understand how cultural experience is translated into language, higher order cognitive behavior, and academic achievement. It is now being assumed that the more variable the environment, and the higher the level of stimulation, the broader the range of possible cognitive and affective experiences. Only when a developing
child is exposed to a rather wide variety of stimuli and only when these stimuli are relatively well matched to the patterns of the child can his cognitive development be maximized.

Hess and Shipman (1965) believe that stimulation during early developmental periods is critical to subsequent growth. They argue that behavior which results in social, educational, and economic poverty is socialized in early childhood. They suggest that the central quality involved in the effects of cultural deprivation is a lack of meaning in the mother-child communication system. In addition they, like Bernstein (1961), hypothesize that the structure of the social system and the structure of the family shape communication and language and that language shapes thought and cognitive styles of problem solving. In the deprived family context this means that the nature of the control system which relates parents to their children restricts the number of and kind of alternatives for action and thought that are available to the child; such constriction precludes a tendency for the child to reflect, to consider and choose among alternatives of speech and action (Hess and Shipman, 1965, p. 870).

The developmental process which may be most severely affected by cultural deprivation is the acquisition and development of language skills. There are some data to suggest that sensory deprivation may not restrict intellectual growth or academic readiness directly, but may severely impede language development, which may adversely affect these other variables, especially the development of reading skills. Much data supports the conclusion that the disadvantaged child has difficulty in school due to a lack of familiarity with the language used by the teacher (Bernstein, 1961; Hess and Shipman, 1965). In fact, books are
often absent from the disadvantaged home and language usage is limited, with gestural and other non-verbal means of communication being very common, with the result that the deprived child experiences insufficient language and reading practice at home, which adversely affects his language development, his reading readiness, and later on his school achievement.

Comparisons of the language and speech of middle and lower class children point to blatant deficits in the lower class group (Templin, 1957; and Irwin, 1948). Templin found that upper class children score higher on an articulation test and suggested that this results from the presence of a better speech model. She also found that upper class children have a significantly larger vocabulary than their lower class counterparts and better word recognition skills. Young (1941) found that higher class children used longer sentences, talked more and were more advanced in the functional use of different speech types than were the disadvantaged children. Many other studies have found similar results and these findings are now commonly accepted by nearly all educators and psychologists.

Brown (1965) found that the disadvantaged child lacked experience in understanding and using relational and positional concepts, past and future tenses, and cause and effect sequences. He concluded that the underprivileged child's vocabulary causes his speech to resemble that of the younger privileged child with consequent effects on his achievement level. John and Goldstein (1964) found that disadvantaged children were limited in their ability to label, discriminate, recognize, categorize, and generalize. Furthermore, they found, as others have, that their opportunity for verbal communication is limited and rarely are these
children read to by adults. In addition, the speech he hears is rather limited and poorly structured, so that a poor model for the acquisition of language and reading skills is provided the disadvantaged child by adult figures within his social milieu.

Long and Henderson (1968, 1971) found that their non-reading disadvantaged school beginners behavioral and conceptual patterns implied a poor self-concept, withdrawal, fear and hostility. They suggest that for these children, before any significant changes to greater academic success can be realized, a warm and close relationship with a tutor or teacher is essential, if the child is to develop the confidence and trust necessary for success in the classroom.

In summary, disadvantaged children's pronunciation and articulation, word variety, sentence length, and use of grammatical and syntactical structures seems retarded when compared to higher class children of the same age. The disadvantaged child often comes from a disrupted home situation where the parents have achieved little academic success themselves. Probably, the most important consequence however of the disadvantaged child's language retardation is the slower and less complete transition from concrete to abstract modes of thought and understanding, which could account for the observed cumulative deficit phenomenon in his academic performance and a permanent educational disability. Often, this poor academic performance results in the development of feelings of frustration, fear, and uselessness which can only lead to disinterest in school, behavior problems, withdrawal, and the development of other defensive, neurotic learning inhibitions.
REMEDIAL PROGRAMS FOR THE DISADVANTAGED UNDERACHIEVING CHILD

To overcome the cumulative pattern of underachievement caused by cultural and sensory deprivation many intervention programs have been developed. These remedial programs have taken numerous forms, viewed the problem from different perspectives, and attacked the syndrome at different points in the developmental process. The most notable of these is of course Head Start. Remedial studies with essentially positive results will be summarized in this section.

Ausubel (1966) suggests that the effects of cultural deprivation on cognitive development are only partly reversible. He argues, like Deutsch and others, that a cumulative deficit pattern exists which can result in permanent retardation if intervention does not occur soon enough. Since the style of the disadvantaged home is perpetuated from generation to generation; early educational intervention is necessary if the cumulative cycle is to be broken. Deutsch (1964) and Hunt (1966) argue that this intervention must begin by age three, since following the first year of the disadvantaged child's life, social and environmental factors progressively impede cognitive development. However, others have argued that remediation can be successfully achieved even though the intervention begins later in the developmental sequence. This argument seems to be especially valid when the targeted cognitive skill is reading or mathematics. Further research is needed to determine the optimal age to begin interventive remedial programs with underachieving disadvantaged children. The trend currently is towards enacting them at the youngest possible age.
Skeels and Dye (1939) removed infants (mean age of 19.4 months) from an institutional setting and placed them in foster homes. The new environments were considered to be much more stimulating than the previous institutional setting. After a period of 2 1/2 years the experimental group showed an average gain of 28.5 IQ points while the contrast group showed an average loss of 26.2 IQ points. In follow-up studies the two groups were found to maintain their divergent patterns into adulthood. These dramatic findings helped reverse the popularity at that time of the genetic predetermined view of intellectual development and led to increased consideration of environmental influences.

Gordon (1969) initiated an intervention program to train parent educators to work with mothers of new-born disadvantaged infants in their own homes. Disadvantaged para-professionals were trained to teach the use and function of a series of stimulation exercises to mothers of deprived children. Significant IQ changes in favor of the experimental group were obtained, which reflected only a maximum of twelve hours of stimulatory instruction to the child. These differences were significant for children at 6 months of age.

Bereiter and Engelmann (1966) developed a program for the disadvantaged which emphasized reasoning and logical thinking as a means of processing information. The children's mean mental age 4 1/2 months after nine months of training showed significant improvement on verbal subtests of the ITPA (Illinois Test of Psycholinguistic Abilities).

Blank and Solomon (1968) developed a program geared towards facilitating the abstract thinking of pre-school deprived children. This special language program employed individual tutoring sessions, on
a daily basis, to effect significant IQ changes within their experimental group. These researchers argued rather convincingly, that exposure to numerous potentially enriching stimuli will not necessarily result in changes in the cognitive deficits of the disadvantaged child. For learning and the desired changes to occur, they hypothesized, the child must actively involve himself with the stimuli and fully comprehend their significance. He must have the ability to organize his thoughts, reflect upon situations, comprehend the meaning of events, and structure his behavior, so as to be able to choose among alternatives of thought and action (Blank and Solomon, 1968, p. 380). The characteristic absence of these abilities in the poorly achieving disadvantaged child; the authors suggest, is the most debilitating deficiency of the culturally deprived child. Consequently they developed a program to aid in the development of abstract language and hypothesized that this would cause these children to develop an internal symbolic system which would subsequently enable them to make more effective use of their experiences.

Strayer (1930), Dawe (1942), Rhinegold (1961), Casler (1965), Sayegh and Dennis (1965), Levenstein (1970), Gray and Klaus (1965), and others have also demonstrated the feasibility of positively altering the effects of early deprivation in young children through various stimulation intervention programs.

Others have attempted to make similar changes later in the developmental process, after the child has entered the public school system. Carter (1967) developed a language stimulation program for disadvantaged children in the first grade. He hoped to alter the language age scores, the mental age, and the reading ability of the children who participated
in his program. He found that the children improved in each area. However, their reading scores were not improved in the short run (pre vs. post), but were facilitated when assessed for long-term impact twenty months after the cessation of the program. This finding lead to the conclusion, as others have suggested, that increased stimulation will result eventually in improved academic performance in socially deprived children.

Shore, et.al. (1971) explored the effectiveness of a Follow-Through program with disadvantaged first grade children. Using a special semi-structured instrument, the Locus of Control Interview, the authors were interested in assessing the changes which resulted from participation in their program in these children's feelings of helplessness and powerlessness, often cited in the literature on the underachieving disadvantaged child as a cause of their poor academic performance. They did find changes in the children's feelings of control over their environment as a consequence of their involvement. However, changes in academic skills were less striking, especially in reading readiness. In addition, these researchers found differences in the effectiveness of the teachers in affecting the desired changes. The teacher who was thought to be the least effective by school administrators was the most effective in changing the children's locus of control and also in bringing about cognitive changes. This teacher was the only one of the three involved who volunteered for the program and was considered by the authors to be the most enthusiastic. She also was the least experienced of the teachers. The authors concluded that the teaching criteria derived from our prior educational experience may not be appropriate and that further study of the characteristics of those who are effective in eliciting changes in
poor achievers certainly is needed.

Graffit (1963) found that the simple addition of forty paperback books to the classroom not only increased significantly the children's achievement levels, but also improved the student's attitudes toward reading. Pfau (1967), obtained changes in attitude, interest, and achievement in his first and second grade readers, simply by adding a planned supplemental program of successful recreational reading to their curriculum. Luser, et al. (1958) employed forty-three phonetic drill sessions to affect changes in reading ability in underprivileged children. Froelich, et al. (1967) developed a successful reading program for disadvantaged children which emphasized and attempted to facilitate the child's oral language development. The program was individualized and highly structured and considered to be a success.

Currently, some educators, parents, and psychologists have been supporting a different approach to the remediation of poor achievement. Rather than removing the child from his regular classroom, and placing him in a special class, some recent data suggests that it is better for him to remain in his classroom and for him to receive specialized attention within that setting.

This attention has taken numerous forms, including intensive counseling and tutoring. Avila, et al. (1968) involved his disadvantaged population in moderately to highly structured skill centered tutoring sessions. The children were tutored by college students for ten sessions during one semester. After tutoring, children exhibited a decrease in their aggressive behavior, an increased score on a leadership index and unchanged scores on a withdrawal index.
Other investigators have employed the disadvantaged underachiever as a teacher or tutor. Gillham (1967) employed twenty eighth graders to work with an overcrowded kindergarten classroom. After the helpers received the commendation of the principal, their peers and parents, and the media, some of the student helpers achieved a gain of as much as four academic years in their reading levels. Morgan and Toy (1970) found similar results. Students who served as tutors showed a mean nine month greater gain on the subtests of the Wide Range Achievement Test (WRAT) than did their non-tutored controls. In addition, the younger tutored children (grades 2-5) made greater gains on the WRAT subtests than did their controls. Haggerty (1970) assigned underachieving adolescent males to one of three groups. The control groups received no treatment. The counseling group met for weekly group counseling sessions. The tutoring group tutored younger students at a local elementary school twice each week. Haggerty found that membership in the tutoring groups produced the greatest achievement and self-concept changes.

Other programs are utilizing specially selected teachers or volunteer tutors, who are aided by academic specialists (reading or mathematics experts, etc.) who devote a scheduled portion of their time to assisting each underachiever and tutor in the content areas where each child is experiencing the greatest difficulty.

Still other approaches have explored the effect of counseling on academic achievement. Proedel, et al. (1960) explored the extent to which group counseling improved the mental health and academic performance of underachieving adolescents. They found that those counseled experienced increased self acceptance, better achievement test scores, and improved
interpersonal skills. Calhoun's (1956) results are essentially in agreement with those of Broedel, et.al. (1960). However, his experimental group did not achieve significant changes in their achievement levels. The changes were in the desired direction however.

Skouksmith and Taylor (1964) examined the effectiveness of non-directive counseling on a group of underachieving twelve and thirteen year old students. They employed two comparable control groups. The twelve children in each group were matched for IQ, age, and previous academic achievement. After six months of counseling there was significantly greater improvement on four of six achievement tests, for the group receiving counseling, when compared to the improvement of the control subjects. In addition, sociogram ratings showed greater improvement in peer and teacher acceptance for those receiving counseling than for either of the control groups.

Baymnr, et.al. (1960), Skouksmith and Taylor (1964), Farwell and Peters (1956), Beurger (1968), and others have further demonstrated that counseling can produce positive changes in children's achievement levels, their self concept and in their interpersonal skills. Unfortunately though, most of these programs have only been enacted with high or average ability underachieving adolescents and their potential effectiveness with underachieving disadvantaged children seems less clear.

Strang (1951) argues that a multitude of pressures all interact to produce a form of maladjustment, which in turn, results in underachievement. Studies by many researchers (Bruner, et.al., 1959; Barrett, 1956; Burns, 1949; Rushton, 1966; Vacc, 1968; Blackham, 1955; Zimmerman, 1965; and Yellott, et.al., 1969) support the conclusion that emotional factors are related to underachievement and
consequently must be considered in any attempt to overcome the problem. Underachieving children, as has been pointed out previously in this review, often feel inadequate and inferior, are immature, emotionally unstable, at times hostile, impulsive, nervous, easily discouraged and often have a poor conception of their own potential and the appropriateness of their own achievement. Consequently, it is easy to see why Rie (1974), Buerger (1968), Farwell and Peters (1956), Cutts and Mosley (1957), and Prentice and Sperry (1964) all recommend the utilization of counseling and therapeutic techniques in the remediation of these problems. They suggest that a therapeutic relationship will help the child overcome his neurotic learning inhibitions. The therapeutic experience they argue will enable the child to confront his anxious expectations and confused impressions about school and therefore gradually allow him to overcome the defensive behaviors and attitudes which have accrued as a result of his continued failure in the classroom. These personal changes should enable the child to overcome his apathy, to feel better about himself, and to develop a level of trust in both his school and his teachers that will allow him to make the fullest use of his abilities. When these therapeutic procedures are coupled with appropriate academic tutoring these theorists suggest that the child will have the opportunity to begin catching up with his peers and eventually the cumulative deficit cycle will be broken.

Since most tutoring takes place outside the classroom, the therapeutic tutor can avoid being the object of the child's anti-teacher feelings and can modify curricular content to suit the skills and needs of the individual child at that particular moment. After a mutual and
interdependent relationship is established between the child and the therapeutic tutor, the interpersonal experiences they enjoy can be extremely rewarding and productive and highly facilitative of desirable personality changes.

Riessman (1968) suggests that large numbers of nonprofessionals should be recruited from the ranks of the poor, trained and then used to help tutor poorly achieving children in the schools. Carkhuff (1969), after reviewing the literature, concludes that extensive evidence exists which indicates that lay persons can be trained to function at minimally facilitative levels. Anker and Walsh (1961), Appleby (1963), Poser (1966), Warme (1965), Brown (1965), Harvey (1964), Guerney (1964), and others have demonstrated that lay persons can effect significant constructive changes with a wide variety of patient types.

Trusty (1971), noting Rogers' (1966) observation that research evidence is accumulating to suggest the conclusion that the interpersonal relationship is the most important factor in facilitating the learning process, developed a therapeutic tutoring reading program. She outlined twelve basic learning principles which she felt were crucial to an effective remedial reading program. After utilizing these principles for five years with many third grade children, Trusty concluded that her approach was more effective than many more traditional ones in improving children's achievement reading scores and also in interesting the children in reading. She found that her children began reading many more books while participating in the program and that this new
found interest seemed to continue long after their involvement in the program had ceased. In fact, she stated, that most of the children report that reading has become their favorite subject. Arthur (1946) also developed a successful therapeutic tutoring program which employed similar principles to Trusty's, and produced similar results. However, there have been remarkably few therapeutic tutoring programs and further work in this area seems warranted before the full effectiveness of this dualistic and reciprocal approach in the remediation of culturally related underachievement will be fully known.

After reviewing this extensive literature, it seems fair to conclude that compensatory education programs with young underachieving children, that have been built on specific learning objectives or environmental features not only report definite gains, but gains that can be related to the particulars of the remedial program. So it appears that intervention during these years can have an impact on the cognitive development and on the achievement level of disadvantaged children who have been ravaged by the affects of cultural and/or sensory deprivation. However, since other studies have found that there is a tendency for these gains to level off after the first year or to dissolve completely with later school experience, the significance of early deprivation as a critical process, difficult to permanently alter in the face of continued nonfunctional environmental deprivation is underlined and reinforced. More research needs to be done to reliably determine the plasticity of the affects of cultural depri-
ration, the long term effectiveness of remedial programs, and the characteristics of the effective counselor, teacher, or tutor.

THE EVIDENCE OF THE OVERALL EFFECTIVENESS OF COUNSELING AND PSYCHOTHERAPY

Buerger (1968), Shore (1971), Torrey (1969), Arthur (1946), Truax and Carkhuff (1967), and many others have recommended that more research be done to determine the dynamics and characteristics of the effective student-tutor relationship. Unfortunately no one has systematically studied the characteristics of the effective therapeutic tutor. Consequently, the related literature on the effectiveness and ineffectiveness of counseling and psychotherapy will be examined in this section, to garner any leads and insights which might prove relevant to this study.

The Evidence of the Ineffectiveness of Counseling and Psychotherapy

Eysenck's (1952, 1955, 1960, 1965) papers stimulated a considerable debate, since they questioned a sacred cow, the value and effectiveness of psychotherapy in producing constructive behavioral and personality changes. He suggested (1960) that psychotherapy be viewed like the wonderous cure developed by Galen, the father of modern medicine. Galen described his treatments thusly: "All who drink this remedy recover in short time, except for those whom it does not help, who all die and have no relief from any other medicine. Therefore, it is obvious that it fails only in incurable cases" (Eysenck, 1960, p. 5). Eysenck's
contention, however, that average counseling and psychotherapy does not result in average client improvement greater than that observed in clients who receive no special counseling or psychotherapeutic treatment does seem to be supported by a number of reasonably sound studies.

Frank (1961) notes that numerous statistical studies have found that two-thirds of their neurotic samples improve immediately after treatment, regardless of the nature of the treatment. However, a similar improvement rate was also found for non-treated equivalent control subjects. Barron and Leary (1955) utilized fully qualified psychoanalytically oriented therapists to treat neurotic patients in a psychiatric clinic. The control group was seen as highly compatible with the treated experimentals on most relevant dimensions (i.e. diagnosis, prognosis, age, sex, etc.). Both experimental groups (one received group and the other individual treatment) evidenced a reduction in pathology as indicated by the neurotic scales of the MMPI (Minnesota Multiphasic Personality Inventory). However, the control group also showed similar reductions in neurotic pathology, resulting in the finding that there were no significant differences between the improvement rates for any of the three groups.

Garfield, et al. (1971) explored the relationship among eight different criteria of psychotherapeutic effectiveness. They found little relationship among the various criteria of change. However, global measures of overall improvement produced the most positive index of desired changes, whereas no significant changes in the behavior or personality of their thirty-four clients were indicated by the other seven outcome measures.
Teuber and Powers (1953) employed a matched pair design with 650 delinquency prone boys assigned randomly to either a treatment or control group. Treatment lasted for as long as eight years. The counselors were of different orientations and there was minimal therapist turnover during the study. The counseling was described by the authors as primarily supportive. Both groups experienced equivalent changes, despite the fact that the therapists and the majority of the boys thought their relationships were effective. In fact, a slight difference indicating greater changes within the control group was found.

Rogers and Dymond (1954) in a methodologically inadequate study with 29 patients who received client-centered counseling found that their treatment procedures resulted in changes in self-concept scores, but not in any noted behavioral changes. This study is criticized however, because the control group did not control for improvement over time, since they were not given an equal opportunity for spontaneous improvement. A follow-up Chicago study by Cartwright and Vogel (1960) employing similar procedures and an adequate control group found essentially no greater improvement in the experimental group than they found spontaneously in their control group.

Brill and Beebe (1955) in an important study of the effectiveness of the treatment of war neuroses during World War II found that therapeutic treatment did not increase significantly the improvement percentages of that group when their change scores were compared with the data from
either control groups; one of which received routine hospital care, rest and sedation; while the other received absolutely no treatment. Thus, in this study as in others, time alone seemed to be the critical factor in affecting positive changes.

Richardson (1960) and Goodstein and Crites (1961) found that counseling did not improve the academic performance of their students when comparisons were made with the academic performance of other non-counselled poorly achieving college students. The latter study in fact found a significant difference in academic achievement favoring their control group.

May and Tuma (1964), in a well designed study, randomly assigned their subjects to either a chemotherapy group, a control group, a psychotherapy plus chemotherapy group, or a psychotherapy only group. Interestingly they found chemotherapy to be generally more effective than psychotherapy.

Frank, et al. (1959) in an important study, controlled for differences between therapists by having each psychiatrist conduct each of the three different treatment modalities utilized in the study. Patients were assigned randomly to either group or individual therapy on a weekly or bi-weekly basis. The control group consisted of those who dropped out of treatment before the fourth session. Each group showed improvement, but again the greatest change was found in the control group. Frank followed twelve of the treated group two to three years after the cessation of therapy and found that many of these patient's symptoms had reappeared
and that treatment was again indicated. This group then received a placebo (an inert medication) and after two weeks showed a significant decrease in their symptoms, which was as great as that which had followed six months of psychotherapy in the earlier study.

Poser (1966) in an interesting study, explored the outcome of psychotherapy by assigning his 343 hospitalized male schizophrenics to one of four groups. One group was treated by experienced psychiatrists, another by social workers and another by untrained college students. The final group served as an untreated control. The findings showed significantly better outcomes for the group treated by untrained college students, suggesting the potential value of untrained or semi-skilled indigenous workers in mental health work.

Levitt (1963) after reviewing twenty-two outcome studies with children found an average overall improvement rate of 62.5%, which was nearly matched by the average overall improvement rate for the controls, who did not receive treatment. It was concluded that psychotherapy does not facilitate recovery from emotional illness in children to any greater degree than no treatment whatever. Powers and Witmer (1951), Mink and Isaacson (1959), Broedel, et al. (1960), Gliedman, et al. (1958), Rogers, et al. (1967), and many others have reached similar conclusions with subjects with different diagnostic classifications and with subjects of different ages.

Shlien (1966) summarized the overall impact of the past two decades of psychotherapeutic research by pointing out, "Continued subscription
(to psychotherapy) is based upon personal conviction, investment, and observation rather than upon general evidence" (Shlien, 1966, p. 125). Certainly a large number of studies involving an equally large number of therapists, suggest that the overall effects of psychotherapeutic interventions is no better than the effects of normal day to day existence (spontaneous recovery).

However, there is some equally strong evidence that some therapy is indeed effective. There are a number of studies, involving certain therapists that do produce positive outcomes. These studies will be discussed in the following section.

Evidence of the Effectiveness of Some Counseling and Psychotherapy

Eysenck's contention (1952, 1960, 1965) that psychotherapy is ineffective sparked considerable controversy among clinicians. Subsequently, his and other studies that have suggested that counseling is at best ineffective have been severely criticized as inaccurate and often as methodologically unsound.

Some of this furor does seem justified. Bergin (1971) re-reviewed the original articles on which Eysenck based his conclusions. The data Eysenck derived from combining these studies are at best ambiguous and inaccurate. In addition different percentages of improvement can be calculated depending upon the criterion and method used by the particular reviewer. Apparently, different investigators with different biases can arrive at quite different conclusions from the same basic data.

There are a number of explanations for the existence of these divergent conclusions about therapeutic outcome, including: a lack of equivalent criteria of outcome, differences in duration and thoroughness
of follow-up studies, a lack of precisely comparable cases between studies, considerable variations in the quantity and quality of the therapy received, variations in the nature, onset, and duration of the subject's disturbance, and often an absence of assessments of outcome made independently of the therapist's own evaluations.

In addition to the limitations outlined above for these earlier studies, recent outcome research and new evidence concerning the level of spontaneous improvement has led Bergin (1971) and others to conclude that on the average, the overall effect of psychotherapy is modestly positive rather than equivalent to no treatment whatever, as Eysenck suggested. Graham (1960) explored the effect of psychoanalytic psychotherapy on sexual enjoyment and satisfaction. He found increased satisfaction and enjoyment as the length of treatment increased. Draspa (1959) examined the effectiveness of psychotherapy in treating individuals with psychosomatic muscular pains. Subjects were assigned to either a routine medical treatment group or a psychotherapy group. The psychotherapy group improved considerably more rapidly than the individual group and at the conclusion of the study had almost twice as many members who were completely free of pain. O'Connor, et al. (1964) published equally as impressive results with another psychosomatic group of patients with ulcerative colitis.

Ends and Page (1957) compared remission rates for patients assigned to analytic and client-centered group therapy with those in leaderless didactic and discussion control groups. They found that the members of the therapeutic groups had remission/non-remission ratios of 1 to 2, while the control group showed a ratio of 1 to 5.
Campbell (1965) followed up a previously studied group of patients, twenty-five years after the cessation of their treatment. After analyzing the 123 male pairs, he discovered that the counseled group showed a statistically significant superiority over the control group on a measure of contribution to society.

Heinichi and Goldman (1960), after reviewing the literature, came to the conclusion that when compared to controls, children who received treatment show a significantly higher percentage of successful adjustments. Seeman, Barry, and Ellenwood's (1964) results were essentially ambiguous. They made a carefully controlled study of eight children, who were receiving play therapy and found a significant change on one pre-post therapy determination. However, their other outcome measures did not change significantly. Not surprisingly, the children's teachers were not able to discern any reliable changes in the children, but then again their sample was extremely small.

Some recent studies have begun to question the spontaneous recovery rates cited by Eysenck (1952, 1960, 1965), Landis (1937), Denker (1947), and others. These early studies found a 72% recovery rate among their untreated neurotic samples and this figure became the baseline with which treated groups were compared for comparisons of that treatment's effectiveness.

Friess and Nelson (1942) found a spontaneous recovery rate of only twenty nine percent. Shore and Massimo (1966) found a rate of thirty percent. Materson (1967) found a thirty-eight percent recovery rate in his adolescent subjects. While Vorster (1966) reported that only thirty-four percent of his neurotic sample had improved spontaneously during a
three year period. Other studies have found essentially similar results with rates ranging from 0 to 70%. However, these studies too are not without their own methodological flaws, and consequently these figures can only be taken as indicative. However, it does appear that Landis', Denker's and Eysenck's rates were overly optimistic and that a truer median spontaneous rate of improvement for neurotic subjects is probably in the order of 30 to 45% (Bergin, 1971).

Several explanations have been put forward to account for the spontaneous improvement noted in the untreated controls that were utilized in the outcome research already reviewed. Some theorists have suggested that these people may seek help from non-professionals, i.e. spouses, friends, teachers, clergy, etc. Indeed, Frank (1961) found that 50% of a group which had sought psychotherapy at his center had also sought help from a variety of non-mental-health sources over a period of several years. If this is indeed the case, then these findings underscore the role that non-psychiatric resources play in determining the mental health of our community and are suggestive of the important role that indigenous workers might play within various psychiatric settings.

Others have argued that personal coping mechanisms and changed environmental conditions could account for spontaneous recovery. These theorists argue that individual self-help actions and defense mechanisms (sublimation, denial, etc.) may result in some improvement, especially for the short term. In addition, these theorists argue that personal philosophies and religions also may play a role in affecting spontaneous improvement. Irregardless, it does appear that the spontaneous remission phenomena supports the conclusion that psychotherapy may be a special form of a broader range of therapeutic phenomena that exist naturally.
within our society. However, its full parameters and dynamics are not yet completely understood.

Numerous studies, many of which were described in another section of this review, have found that counseling has a facilitative effect on the academic achievement of poor and/or underachievers. Likewise, many behavior modification studies have reported favorable outcomes (i.e. Lazurus, 1961; Lang, 1965; Bandura, 1969; Wolpe, 1958; etc.; etc.).

Cross (1964) reviewed nine controlled studies and concluded that six reported favorable outcomes resulting from psychotherapy. Dittman (1966) added five more studies to Cross's group and decided that ten out of fourteen controlled outcome studies are favorable to psychotherapy. Bergin (1967) however feels that these reports are overly optimistic, inaccurate, and misleading and he goes on to argue that overall, average psychotherapy, at best, results in only modest improvement.

Meltzoff (1969) and Meltzoff and Koonreich (1970) disagree with Bergin's conclusion. They argue that over 100 studies support the conclusion that the effectiveness of psychotherapy, with a wide variety of patient types, as ordinarily performed by journeymen therapists, has been clearly documented. The studies upon which their conclusions are based include many of those already mentioned and Persons (1967), Feifel and Ellis (1963), Betz (1962), Ellis (1956), Vorster (1966), Orgel (1958), Shore and Massimo (1966), Riess (1967), Uhlenhuth and Duncan (1968), and others too numerous to mention here.

Kellner's reviews (1965 and 1967) of the extensive outcome literature lead him to reach more cautious conclusions on the effectiveness of psychotherapy than those reached by Meltzoff. His position is essentially in agreement with Bergin (1971). Kellner argues that the results of out-
come studies are frequently in conflict, that comparability of studies is often not possible, and that often divergent processes are occurring in the diverse influences receiving the general label psychotherapy. He suggests that the more homogeneous the intervention and the sample of patients, and the more specific the criterion, the more favorable the outcome results. Finally, he concludes that psychotherapy is effective under restricted, specific circumstances, but generally outcome studies, though sometimes positive, produce mostly ambiguous results (Kellner, 1965 and 1967).

Thus, from the evidence currently available, it can be concluded that overall, average psychotherapy has only modest positive effects. However, since a number of studies (Bergin, 1966 and 1967; Truax, 1963; Eysenck, 1960; Truax and Carkhuff, 1964 and 1967; Cartwright and Vogel, 1960; etc.) have found a significant increase in the variability of criterion scores for the treatment groups at post-testing, it appears that some treatment cases are improving while others are deteriorating. Thus, some counselors and psychotherapists are effective, while others are destructive. Then, when the positive and negative effects are averaged together within the experimental group, they cancel each other out to some extent, and result in an overall improvement that is equal to or only slightly greater than the change occurring within the control group due to spontaneous remission.

Truax, et al.'s (1966) data demonstrate this beautifully. With their forty outpatients at John Hopkins, who received individual psychotherapy from resident psychiatrists, they found an overall improvement rate of 70% (nearly the level of spontaneous recovery Eysenck proposed for untreated populations serving as controls). However, those who saw
therapists seen as effective by the authors had an improvement rate of 90%, while those seeing the poorer therapists had an improvement rate of only 50%, well below Eysenck's possibly inflated rate.

It appears that the most important conclusion that can be drawn from the outcome studies reviewed here is that it is impossible to conclude very much from gross studies of therapeutic effects. Only by breaking therapy down into its primary components will we be able to garner a full understanding of this complex process. "We need to determine what treatment, by whom, is most effective for this individual, with this specific problem, under this set of circumstances" (Paul, 1967, p. 111). Thus, we must explore the therapeutic process itself if we are to isolate the effective elements within the process that lead or result in constructive changes within the patient.

THE PRIMARY INGREDIENTS OF EFFECTIVE COUNSELING AND PSYCHOTHERAPY: SOME THEORETICAL VIEWS

Clinicians today are certainly aware of the many competing theories and schools of counseling and psychotherapy. The inability of any one group or position to prove itself universally correct in both prediction and practice has produced a prolific situation where several new approaches are proposed each year.

Thorne (1950) attacked the then current therapeutic sophistication and routine clinical procedures. He criticized the therapeutic community in which large numbers of competing schools of psychotherapy are widely applied and where almost no validation and standardization procedures are employed. In addition, Thorne hypothesized that since
many methods of treatment appear to produce favorable results, it seems possible that similar factors must be operating among these seemingly diverse methods. If this is indeed the case, therapists must be concerned with discerning the common elements among all the effective psychotherapeutic systems and procedures. Because of the numerous theories of psychotherapy, this review will try to ignore the individual theories concentrating instead on reviewing those determinants of change that others have considered to be common to a number of effective therapies.

Black (1952) eloquently emphasized the therapist-client relationship and the therapist's interpersonal skills, which he considers to be the central therapeutic ingredients. He writes, "our lack of progress in understanding the process of therapy and in developing a systematic methodology may lie partly in our preoccupation with specific techniques and schools, to the neglect of what may well account for the most significant share of the behavioral changes produced by psychotherapy - the interpersonal relationship itself" (Black, 1952, p. 303). Rogers (1951, 1957) goes even further by contending that the relationship is therapy. Ruesch and Prestwood (1950) buttress Roger's contention, by suggesting that the effective therapeutic agent(s) are found within the interaction between the patient and his therapist.

Hobbs (1968) outlined several possible sources of gain in psychotherapy. These included: 1) the therapeutic relationship which is intimate and safe, 2) the transference relationship, 3) the divestment of verbal and other symbols of their anxiety provoking potential, and 4) as a result of the therapist placing the locus of control on the client, cognitive structures develop which enable the client to feel that he can better control his destiny.
Rogers (1957) in a classic work, listed six necessary and sufficient conditions for positive therapeutic change (change in personality structure in the direction of greater integration and less internal conflict). The six conditions he proposed are:

1) that the counselor and the client are in psychological contact
2) that the client is vulnerable
3) that the counselor is congruent and integrated in the relationship
4) that the counselor has unconditional positive regard for the client
5) that the counselor empathizes with the client, and
6) that the empathic understanding is, to at least a minimal degree, communicated to the client (Rogers, 1957).

Despite the multitude of theories of psychotherapy and the differences in terminology found within these theories, a number of common ideas can be extracted from the principle approaches: analytic, client-centered, behavioristic, eclectic. Most have stressed the importance of a therapist's ability to be integrated, mature, genuine, and authentic or congruent in his relationship with his patient. In addition most have stressed the importance of the therapist's ability to provide a trusting, safe, nonthreatening atmosphere through his acceptance, warmth, unconditional regard, or love. Finally, nearly all theories of psychotherapy suggest that if the therapist is to be fully effective, he must be with the client and capable of totally understanding him (Truax and Carkhuff, 1967, p. 25).

Truax and Carkhuff (1967) termed these three characteristics: accurate empathy, genuineness, and non-possessive warmth and suggested that they were the universal characteristics of all effective therapy,
regardless of the therapist's orientation, personality, and/or patient population. However, these ideas were not new. In fact, many of these same concepts can be found, in different words, in the early psychoanalytic writings of Freud, Rank, Allen, Taft, and others.

Adler placed considerable emphasis on understanding and he considered therapist warmth to be central to any effective therapy. In addition Adler stresses the importance of love, appreciation, and collaboration in any effective therapeutic interaction (Hall and Lindzey, 1957). Shoben (1953), a learning theorist, considers empathic understanding and warmth to be important in psychotherapy. Eclectic theorists Noyes and Kolb (1964) and Brammer and Shostrom (1964) stress the therapeutic triad of empathy, warmth, and genuineness as basic attributes of an effective counselor. Bordin (1955), Ferenczi (1930), Alexander (1948), Schaefer (1959), Strupp (1960), Halpern and Lesser (1960), Hobbs (1962), Dymond (1949), Truax (1961), and others too numerous to mention here have all emphasized the importance of the therapist's ability to act genuinely and to understand sensitively and accurately the inner experiences of the patient. Certainly, there is considerable agreement among these theorists of the important role played by these three therapeutic conditions in all effective psychotherapies. The common agreement among leading therapists of diverse orientations of the central characteristics of effective therapists was nicely demonstrated in a study done by Fiedler (1950, 1951). He found that experienced therapists of different schools of psychotherapy agreed on the elements of an ideal therapeutic relationship, which they characterized as warmth, acceptance, and understanding.
Rogers (1967) argued that due to the reciprocal nature of the triad of central therapeutic ingredients, the order in which one considers them, becomes important. Truax and Carkhuff (1967) also point this out. "To be facilitative toward another human being requires that we be deeply sensitive to his moment to moment experience, grasping both the core meaning and significance and the content of his experiences and feelings. Such deep empathic understanding requires first that we have at least a degree of warmth and respect for the other person. Thus, empathic understanding can scarcely exist without a prior or concomitant feeling of nonpossessive warmth. In turn, neither the empathy nor the warmth could be constructively meaningful in any human encounter unless it were real. Unless the counselor or therapist is genuine in relating to the client, his warmth and empathy may even have a potentially threatening meaning." (Truax and Carkhuff, 1967, p. 32). Thus, genuineness or authenticity is seen by these theorists as basic to an effective relationship. After the reality of the therapist is established, the warmth and respect communicated to the patient becomes another integral component of the relationship. Then finally, given a relationship that is warm and genuine, the therapeutic process proceeds because of the therapists empathic understanding of the meaning of the client's actions and feelings.

However, it should be emphasized that there will still be considerable variation in the degree of change observed in clients seeing high or low condition therapists. This may be accounted for by client variation and differences and/or variation in the therapist's reinforcement pattern,
modeling level, etc., etc. We just do not yet know. Truax and Mitchell (1971) suggest that the three therapeutic conditions affect a patient's level of change in four different ways: "(1) they serve to reinforce positive aspects of the patient's self-concept, modifying the existing concepts and thereby leading to changes in the patient's own self-reinforcement system; (2) they serve to reinforce self-exploratory behavior and thereby elicit self concepts and anxiety laden material that can then be modified by selective reinforcement; (3) they serve to extinguish anxiety or fear responses associated with specific cues, both those elicited by the relationship with the therapist and those elicited by patient self-exploration; and (4) they serve to reinforce human relating, encountering, or interacting, and serve to extinguish fear or avoidance responses associated with human relating" (Truax and Mitchell, 1971, p.322).

Truax (1965) explored the possibility that the therapist may selectively offer empathy, warmth, etc. contingent on in therapy verbalizations and behaviors. He studied a single-long term case treated by Carl Rogers. He found that indeed the therapist did respond selectively with differential levels of the three conditions for five of nine classes of patient behavior. If this finding is found to be valid in a wider segment of therapeutic interactions, its random application may account for, at least in part, the differential effectiveness of a therapist with different clients. In addition, Truax found that the selective reinforcement by the therapist of these five classes resulted in an increase in frequency for four of the five. Further research will be needed to determine what other factors affect outcome. Now, however, let us turn our attention to reviewing the specific therapeutic conditions themselves.
The Meaning of Therapeutic Genuineness

Everyone has encountered individuals they consider to be artificial, insincere, or fakey. These people operate behind a front and say and do things for an impression, because it sounds good rather than because they really believe in what they are doing. Consequently, we tend to distrust them and usually relate better to people we see as open, genuine, and trustworthy.

Rogers' (1967) idea of therapist congruence conveys these ideas and is very similar to Truax and Carkhuff's concept of genuineness. Rogers believes that the therapeutic relationship, and indeed all effective interpersonal relationships, must exist without defensiveness and retreats into facades or roles. The effective therapist must be honest with the client and with himself. He must be aware of himself, in that the feelings he is experiencing are available to him and to his awareness, and he must be able to live these feelings, to be them in the relationship, and to communicate them if appropriate. He is being himself, not denying himself (Rogers, 1967).

Frieda Fromm-Reichmann (1952) also stresses this lack of defensiveness in an effective therapist. Rotter (1964) proposes that a therapist should be accepting, sympathetic and interested in his patient and free from interfering neurotic problems of his own, which may result in an undesirable countertransference reaction to the patient. Moustakas (1959) and the existential theorists have further emphasized the importance of genuineness in effective therapeutic encounters.

In summary, the genuine therapist is open to all types of feelings and experiences, both positive and negative, without defensiveness retreats into intellectualization and/or professionalism. He is aware of
himself and his participation in the clinical encounter and he is able to recognize and accept his own strengths and weaknesses as they relate to each therapeutic interaction.

**The Meaning of Therapeutic Warmth**

Rogers (1951, 1957, 1967) has emphasized the nonpossessive and unconditional warmth found in an effective therapeutic relationship. "The therapist communicates to his client a deep and genuine caring for him, as a person, with human potentialities, a caring uncontaminated by evaluations of his thoughts, feelings, or behaviors. The therapist experiences a warm acceptance of the client's experience as being a part of the client as a person, and places no conditions on his acceptance and warmth. He prizes the client in a total rather than a conditional way. He does not accept certain feelings in the client and disapprove others. He feels an unconditional positive regard or warmth for this person. It involves as much a feeling of acceptance for the client's expression of painful, hostile, defensive or abnormal feelings as for his expression of good, positive mature feelings" (Rogers, 1967, pp. 102-103).

Rotter (1964), a social learning theorist, describes the importance of therapist warmth within behavior modification therapies. "...it is necessary that the patient trust him (the therapist) and accept his objectivity in the situation. Consequently, the good therapist is 'warm' and communicates to his patient his concern and interest in him" (Rotter, 1964, p. 86).

Cameron (1963) writes of therapeutic warmth, "the patient must feel free to establish psychological closeness with the therapist, or to put psychological distance between himself and his therapist, without fear
of arousing personal affection or personal offense. The therapist, for his part, remains always friendly, warm, firm, and accepting, but he avoids anything that approaches a personal entanglement with his patient. The competent psychotherapist is not offended by a patient's coolness, criticism, hostility, or provocation" (Cameron, 1963, p. 769).

White (1948) considers warmth to be a central ingredient of all effective forms of psychotherapy. He emphasizes however that this warmth must be genuine. "The therapist is an expert... the therapist is permissive, the therapist is interested and friendly, communicating in this way a certain warmth that makes the relationship more personal than is ordinarily the case" (White, 1948, p. 45).

The Meaning of Accurate Empathy

The therapist's ability to accurately and sensitively understand the client's experiences and feelings and their particular meaning to him, stems from his genuineness and warmth. An accurately empathic therapist is fully in touch with his client and completely at home within his universe. Rogers (1967) captured its essence. "It is a moment to moment sensitivity that is in the here and now, the immediate present. It is a sensing of the client's inner world of private personal meanings as if it were the therapist's own, but without ever losing the 'as if' quality" (Rogers, 1967, p. 104).

Others have similar views of empathy. Norman and Ainsworth (1954) describe it as the ability to put one's self in another's place. One must realistically view another individual. Fiedler (1953) sees empathy as the ability to perceive another person. Hogan (1969) views empathy as a tendency to modify one's behavior as a result of another's feelings or behavior. English and English (1956) see empathy as a cognitive
process. They suggest it is an apprehension of the state of mind without any associated feeling (as in sympathy) of what the client himself feels.

However, the ability and sensitivity to communicate this understanding back to the client is also a major component of accurate empathy. The empathic therapist must not only sense the affective state of the client accurately, but he must also communicate this perception in a language that will aid the client's self perception and which the client can accept and understand. The empathic therapist is with the client and his responses reflect their synchronization. The therapist who is very empathic will indicate not only a sensitive understanding of the apparent feelings but will by his communication clarify and expand the patient's awareness of these feelings or experiences.

Burnham (1961) stresses the idea that empathic understanding does not imply a passive phenomena, whereby the therapist somehow slowly begins to understand the patient and his dreams, thoughts, and fantasies. Rather, he observes how the client behaves and gradually formulates hypotheses as to why he behaves in a particular way. In order to accomplish this however he must first gain some understanding of how the world looks to the patient. "Therefore, the therapist tries by observation, inference, and guess work to gain a reasonably accurate picture of the flow of thoughts and feelings across the stage of a patient's consciousness, and to grasp the representation of reality which the patient organizes from this flow of conscious experience" (Burnham, 1961, p.210).

Ginott (1965) has emphasized the importance of accurate empathic understanding in normal child development. "How can we help a child to know his feelings? We can do so by serving as a mirror to his emotions. A child learns about his physical likeness by seeing his image in a
mirror. He learns about his emotional likeness by hearing his feelings reflected by us " (Ginott, 1965, p. 35).

In summary, the accurately empathic therapist is with his client. The therapist's responses move, with sensitivity and accuracy, into feelings and experiences the client may only partially understand. Underlying feelings or experiences are specifically identified and gradually become understood. The therapist's words and actions reflect a togetherness with the patient and his interpretations reflect his understanding and can be accepted, without defensiveness, by the client. Finally, it has become clear, that as a result of the therapist's warmth and genuineness, his sensitivity and empathic understanding, and his communication of the patient's inner state, the client is encouraged and self-exploration and personality growth are the inevitable result.

The Development of a System to Assess Accurate Empathy, Warmth, and Genuineness in Counselors and Therapists

After realizing that this triad of central therapeutic ingredients was so essential in effective counseling and psychotherapy, researchers at the University of Wisconsin attempted to develop a scale to measure each dimension. Initially the scales were very closely tied to Roger's (1957) work and were heavily dependent on the raters' own biases and expertise. However, as the evidence accumulated that empathy and unconditional positive regard were not closely related to positive therapeutic outcomes, interest began to shift to two of the elements already discussed: accurate empathy and nonpossessive warmth. Also, Rogers idea of congruence became less popular and was replaced by genuineness, the absence of defensiveness and phoniness in the therapist.
Truax developed the Accurate Empathy Scale in 1961 and the Unconditional Positive Regard and Genuineness Scales in 1962. The scales were designed for use with live observations or tape recordings of counseling or therapy interviews. They were intended for individual counseling though they have been employed with groups. Although the author considers these scales to be only a beginning attempt at specifying the operational meaning of these concepts and consequently rather inferential and crude, they have not been revised since 1962.

These scales have been widely used in a number of studies and this literature will be reviewed in the following section. Regardless of whether an assessment is made of counseling or therapy or group or individual sessions, the interjudge reliabilities of ratings on the scales for the same samples of therapeutic transactions are rather high. Truax and Carkhuff (1967) reported reliability data from a number of studies. They found a mean interjudge reliability (Pearson's or Ebel interclass) of .739 (averaged from 29 different determinations in 24 studies) with a range of .43 to .95 for the Accurate Empathy Scale. The mean interjudge reliability (Pearson's or Ebel) for the Nonpossessive Warmth Scale was .705 (averaged from 24 different figures from 20 studies) with a range of .48 to .95. The figures were nearly as good for the Genuineness Scale. Its mean interjudge reliability (Pearson's or Ebel) was .621 (averaged from 20 determinations collected from 17 studies) with a range of .25 to .95. Rate rerate reliability over varying time frames have also been found to be consistently high. The validity of these scales is naturally more difficult to assess, but the scales do seem to be related to a variety of therapeutic outcomes.
The Accurate Empathy Scale (See Appendix A) consists of nine different stages or levels of empathic understanding each being operationally defined (This study will use Bergin and Solomon's 1970 ten-point modification of Truax's original accurate empathy scale. The modified scale contains an additional level between the old stage 2 and 3. This modified scale has been utilized in many recent studies). The focus of the scale is on the communication of empathy rather than on attitude or intention. "At a high level of empathy the therapist's remarks fit perfectly with the client's mood and content. His responses not only indicate his sensitive understanding of the obvious feelings, but also serve to clarify and expand the client's own awareness of his own feelings or experiences. At a low level of accurate empathy the therapist may go off on a tangent of his own or may misinterpret what the patient is feeling. At a very low level he may be so pre-occupied and interested in his own intellectual interpretations that he is scarcely aware of the client's being" (Truax and Carkhuff, 1967, p. 46).

The Nonpossessive Warmth Scale (See Appendix B) is similar except that it consists merely of five stages or levels. The dimension ranges from a high level where the therapist warmly accepts the patient's experience without conditions; to a low level where the therapist evaluates a patient and his feelings, and expresses dislike, disapproval, or warmth in a selective and evaluative way (Truax and Carkhuff, 1967).
The Therapist Genuineness Scale (see Appendix C) also consists of five stages ranging from a low where the therapist presents a facade and/or defends or denies his feelings to a high level where the therapist is truly comfortable, open and honest with the client (Truax and Carkhuff, 1967).

In summary, Rogers theorizes that the primary role of nonpossessive warmth is to preserve the client's ego and self-respect and to create a safe and trusting environment for him. The purpose of genuineness is to produce an open and honest non-defensive relationship that will allow for the open discussion of unpleasant truths and finally, the function of accurate empathy is to further the therapeutic process itself.

THE PRIMARY INGREDIENTS OF EFFECTIVE COUNSELING AND PSYCHOTHERAPY: THE EMPIRICAL DATA

The triad of central therapeutic ingredients described in the last section of this review derived primarily from the work of Carl Rogers and his associates. Rogers (1957) went further than most theorists in suggesting that accurate empathy, warmth, and genuineness (empathic understanding, unconditional positive regard, and congruence) were the necessary and sufficient conditions for desirable therapeutic outcome. Others considered this to be folly and argued that these three therapist characteristics could never account for therapeutic outcome. Rogers suggestions did however stimulate considerable research aimed at determining the characteristics of effective therapists. Some of this literature will be reviewed in this section.

Psychotherapy is an extremely complex transactional or interactional process, which at first glance might seem very difficult to study scien-
ically. Effective research would seemingly require continuous monitoring, review, and feedback of the interactive process. Considerable data on a large number of variables would accrue and the tremendous variability of both the therapists and the clients would only seem to exacerbate the outlook. However, despite these problems, as has already been mentioned, a number of important therapist qualities have been identified which are central to most theoretical models of therapy and considered to be necessary for desired client changes to occur.

Whitehorn's and Betz's classic work at Johns Hopkins Hospital (Betz, 1963; Whitehorn, 1964; Whitehorn and Betz, 1957) demonstrated clearly the differential effectiveness of some therapists. Their retrospective study compared a group of seven psychiatrists who had an improvement rate of 75% with their schizophrenic patients with another group of psychiatrists, who had similar training and an improvement rate of only 27%. They discovered that the successful therapists were warm and attempted to understand the patient in a personal, immediate and empathic way. The less successful therapists related more impersonally and focused on the particular psychopathology and an external kind of understanding. The patient groups were seen as very similar.

Halkides (1958) in an early study, selected brief samples from early and late therapy interviews for ten highly successful and also from ten unsuccessful cases. Ratings were made, using a brief scale based on Rogers (1957) work, of the therapist's level of empathic understanding, unconditional positive regard, and self-congruence. She found that successful therapists score significantly higher on these three variables. Hart (1960) however was unable to replicate her findings, possibly because of the
global and vague nature of the scales used and also possibly because he used clinically less experienced raters.

Truax (1961a) found that the presence of these three central ingredients in group psychotherapy with hospitalized patients was highly related to the patient's level of involvement in the therapeutic process, his self-evaluation, and his self-exploration. In another study Truax (1961b) found that effective therapists were rated significantly higher on an Accurate Empathy Scale. Three hundred eighty-four, two minute samples from the middle third of the therapy sessions, were selected. Interestingly Truax found that the therapists level of accurate empathy did not vary significantly during the entire six months of psychotherapy.

Truax (1963) essentially replicated some of his earlier findings. He selected a four minute, tape recorded section from every fifth interview with 14 schizophrenic patients. Undergraduate students were trained to be raters (until they achieved a rate-rerate and inter-rater reliability of at least $r = .50$) so as to avoid any theoretical biases of more experienced raters. The ratings were scored blind. The correlation between the level of the therapists accurate empathy and outcome (both test and diagnostic data) was high ($r = .77$, $p < .01$). Those patients who improved had therapist's rated consistently higher on accurate empathy, while those patients who deteriorated did not have highly empathic therapists. High correlations with outcome were also attained on the Nonpossessive Warmth and for the Therapist Genuineness Scale, even though separate raters were employed for each scale. Control patients showed only moderate gains. Intercorrelations between the three conditions were high (between .25 and .54) demonstrating that be-
tween 6 and 30 percent of the variance is common to each measure. Importantly the positive relationship between accurate empathy and outcome was found for both outpatients and hospitalized schizophrenics. In general, those who received high levels of the three conditions spent more time out of the hospital than those who received lower levels.

In summarizing the results of the four year study of psychotherapy with 16 hospitalized schizophrenics done at the University of Wisconsin, Truax and Mitchell (1971) write that the findings gathered from several studies indicate that: "(a) patients receiving psychotherapy and those receiving control conditions showed little difference in average constructive personality change and particularly no difference in subsequent hospitalization, but that (b) patients whose therapists offered high levels of nonpossessive warmth, genuineness, and accurate empathic understanding showed significant positive personality and behavior change on a wide variety of indices and (c) patients whose therapists offered relatively low levels of those interpersonal skills during therapy exhibited deterioration in personality and behavioral functioning" (Truax and Mitchell, 1971, p. 302). Truax's (1970) findings demonstrated the long term effects of therapists empathy, warmth, and genuineness. He did a follow-up study of the patients studied by Rogers and his associates at Wisconsin. The subject's hospital records were examined for nine years prior to therapy and followed-up for nine years post-therapy. The findings suggested that those seeing effective therapists were released sooner and tended to stay out of the hospital longer. Patients who saw ineffective therapists tended not to be released from the hospital and had a higher return record if they were released. The authors
concluded that an effective therapeutic relationship can exert an influence even over a nine year period.

Truax and Wargo (1967a) studied 160 hospitalized patients and found significantly greater improvement in those who saw empathic, warm, and genuine therapists. In another study which employed analysis of covariance to control for any initial differences in their eighty delinquents, Truax and Wargo (1967b) found a similar though somewhat stronger association between the presence of the three therapist characteristics and positive outcome from group psychotherapy. A later study, Truax and Wargo (1969) replicated the earlier design and obtained essentially similar findings overall. However, when accurate empathy, nonpossessive warmth, and genuineness were analyzed separately, the data suggested nonpossessive warmth and genuineness, though to a lesser degree, were more critical than accurate empathy in effecting changes in outpatients receiving group psychotherapy.

Truax, Wargo, and Carkhuff's (1966) findings were essentially in agreement. They followed up eighty outpatients who received group psychotherapy in an earlier study and found nonpossessive warmth to be especially important for outcome with outpatients receiving group psychotherapy. Genuineness and empathy, though less important, were still related to the outcome. Overall, those who received high levels on all conditions showed above average improvement on 21 of the 23 measures of outcome (p < .001).
However, Truax, Carkhuff and Kodman (1965) found in another study with forty chronic hospitalized patients receiving group therapy, that not all three of the central ingredients were positively correlated with their outcome measures. Genuineness was found to be negatively correlated with warmth and empathy and consequently with the therapeutic outcome. Another study done by Truax, et al. (1966a) at John Hopkins further supported the conclusion that when two conditions of the therapeutic triad are highly related, but the third is negatively related, the prediction of outcome should be based on the two that are most highly related. In this study warmth was negatively related to empathy and genuineness. In general, this finding suggests that when any two of the three central therapeutic conditions are sufficiently high, positive patient or client change will result. This puzzling finding may be accounted for by the interrelationships between the three conditions themselves, in that particular sample of therapists: the ranking of therapists on empathy and genuineness was identical ($r = 1.00$), while nonpossessive warmth was negatively related to both of the other conditions ($r = -0.40$). In both studies (genuineness for the earlier study and warmth for the Hopkins study) the therapeutic condition that was negatively related to the other two was not itself absolutely low for any group of patients. It seems possible, as Gendlin and Geist (1962) propose, that when any one of the therapeutic conditions are sufficiently low it will interfere with the outcome, regardless of the level of the
other two conditions. Thus, in effect the measurement of genuineness alone may serve only to indicate the deleterious effect when its level is low. These findings, overall, however do give support to the conclusion that therapist empathy and warmth are related to positive outcome. These conclusions are valid for group and individual psychotherapy although the magnitude of the association between the two ingredients and patient change in group psychotherapy was somewhat less than that obtained from several studies of individual psychotherapy. The data on the importance of therapist genuineness in group therapy is however in sharp contrast to the findings from individual psychotherapy.

Truax (1962a) was interested in determining whether the average level of accurate empathy is the important factor in affecting change or whether a greater impact results from those less frequent moments of high empathic understanding. To examine this question he studied complete interviews from early and late therapy sessions for 14 schizophrenic patients. Consecutive five minute samples were made throughout the twenty-eight interviews and since the therapists' level of accurate empathy varied from moment to moment, an analysis was made for high and low moments and also for the mean levels overall. The findings indicated that although patients receiving average levels of higher empathic ratings were those who showed improvement, the highest moments of accurate empathy obtained throughout the interviews were more predictive of outcome when compared with cases with relatively lower highest moments. In addition, these
Low moments of empathy were found not to be predictive of the outcome.

Cartwright and Lerner (1963), using a therapist-patient discrepancy measure of empathic understanding, found that the final, but not the initial level of empathy, was related to patient improvement. However, Truax and Carkhuff (1963), Melloh (1964), and Kiesler, et al. (1957) found that their therapist's level of empathic understanding did not tend to vary systematically across time, especially after the eighth session. Bergin and Jasper (1969) found a moderate inverse relationship between therapist MMPI disturbance levels and the degree of empathic understanding the therapist showed during interviews. In addition, they found no correlation between the therapists intellectual ability and his academic achievement and their level of empathic understanding of their patients.

In a related study, Dickenson and Truax (1966) explored the impact of time limited group counseling on the achievement levels of underachieving emotionally disturbed college students. They employed matched therapy and control groups consisting of 48 students. Their findings indicated that the treated group showed significant improvement when compared with their controls. When intra and inter group comparisons were made, students who received relatively high levels of the three therapeutic conditions were found to have made the greatest improvement. However, only this group's improvement exceeded the improvement evidenced by the control group. Thus students who received only moderate levels of accurate empathy, warmth and genuineness did not differ in their achievement levels from those found for the controls.
Stoffer (1968) demonstrated nicely that untrained people can be helpful and empathic. He employed adult female volunteers to work with poorly achieving behaviorally disturbed children. The volunteers were given no training or supervision but were instructed to establish good relations with their child. They met with the child twice a week during a three month period. The volunteers averaged 38 years of age with a mean educational attainment of 13.3 years. The thirty-five children involved ranged in age from 7 to 12 years and all were referred by school psychologists because of behavioral and academic problems. Taped ratings were made of the interaction between the volunteer and her child. High levels of nonpossessive warmth were found to be significantly related to gains in the child's achievement level and in reductions in teacher ratings of his level of behavior problems. High ratings of accurate empathy were also found to be related to achievement gains for ratings derived from late interview sessions.

Truax and his associates have been interested in determining whether these three central therapeutic skills can be taught to untrained inexperienced counselors in a relatively short period of time. Several studies (reviewed in Truax and Carkhuff, 1967) show that graduate students and lay trainees show closely similar levels of therapeutic conditions after only 100 hours of training to the level of these conditions offered by experienced counselors. Truax and Silber (1966) found that only thirty-five hours of training resulted in significant gains in accurate empathy and warmth ratings for their graduate trainees and they concluded that appropriate training can improve a counselor's effectiveness by increasing his ability to be genuine, warm, and empathic with his clients.
Garfield and Bergin (1971) however question the applicability and necessity of the three therapeutic conditions in non-client-centered psychotherapy. In studying the relationship of warmth, genuineness, and accurate empathy with outcome with therapists who were predominately not client-centered, they found no relationship between the three therapeutic conditions and the several measures of outcome they employed.

However, Bergin and Solomon (1963), Spotts (1962), Strupp, et al. (1963), Lesser (1961), Wargo (1962), Truax, Wargo, and Silber (1966) Wyrick and Mitchell (1969), Cartwright and Lerner (1963), and many others have found one or more of the three central therapeutic conditions to be related to positive therapeutic gains with different patient groups, therapists of different orientations, and with different types or modes of treatment. Consequently, it seems safe to conclude that the same therapeutic conditions are indeed effective for a wide variety of human beings, regardless of their particular diagnosis or psychopathology. Therapists who are accurately empathic, genuine, and nonpossessively warm are effective regardless of their particular theoretical orientation. In addition, it seems safe to conclude that the evidence suggests that low levels of these three therapeutic conditions may be related to the deterioration of a patient in therapy.

Once it was demonstrated that these three conditions were related to positive therapeutic outcome the question of causation became important. Is it the client, the therapist, or both who determine the level of a particular condition that will occur during the interactions? The question is an interesting one and unfortunately the relative contribution of the client and the therapist to the core facilitative interpersonal skills is not yet completely known. Currently, the findings
are somewhat contradictory, though the vast majority of the evidence seems to support the conclusion that the therapist is the primary contributor to the levels he offers his clients (Truax, 1963; Truax, 1962a; Truax, et al., 1966a; Anderson, 1968; Alexik and Carkhuff, 1967; Truax and Carkhuff, 1967; and Hirachberg, Carkhuff, and Berenson, 1966; etc.).

Truax and Carkhuff's (1966) study demonstrated this contention rather nicely. Twenty four patients were seen by eight different therapists. Samples were selected so that each of the eight therapists had treated the same eight patients. Analysis of the ratings indicated that the therapists tended to provide different levels of the core conditions when interacting with exactly the same patients. However, different patients did not tend to receive different levels of the therapeutic conditions when interacting with the same therapists. Thus, they concluded that the therapist was the primary determiner of the levels of therapeutic conditions present during any therapeutic encounter.

Other studies have explored other variables that might be related to the level of therapeutic conditions present in any given therapeutic process. Bergin and Solomon (1964) explored the hypothesis that therapist personality disturbance would interfere with their therapeutic efficacy, especially their empathic responsivity. Their findings lent support to this hypothesis. Other factors that may influence the level of therapeutic conditions offered are patient differences such as severity of disturbance (Rogers, Gendlin, Kiesler and Truax, 1967), diagnostic type (Truax and Carkhuff, 1967) and therapist pre-session mood (Gurman, 1973a).

It appears that an overwhelming literature exists supporting the
contention that a therapist or any helping person will be more effective if he is empathic, genuine, and warm. These findings have obvious and important implications for our clinical training programs, our research, and indeed all our interpersonal relationships, both professional and personal. It has become clear, from converging research data, that the human encounter itself, even when intended to be helpful, can be for better or worse, whether we are focusing on behavioral, personality, or academic functioning. It is our responsibility as professionals to guarantee the most favorable outcome that continued research, training, and hard work can insure. The possibilities are endless and the challenge stimulating and rewarding.

THE ASSESSMENT OF ACCURATE EMPATHY, NONPOSSESSIVE WARMTH AND GENUINENESS: METHODOLOGICAL CONSIDERATIONS

To assess the level of accurate empathy, warmth, and genuineness present during any therapeutic interaction, trained raters or judges have applied the Truax scales to samples of counseling or therapy sessions. Often, brief samples of three or four minutes are excerpted from audio-tape recordings of the sessions. Usually one sample is taken from the middle of the interview and one sample from the middle of the final third of the session. However, other studies have varied the procedure somewhat, taking a segment from each third of the session; while others have evaluated, apparently accurately, samples as long as 16 minutes or even a half-hour or as short as a single minute. Still others have evaluated a specified number of therapist-patient-therapist dialogues, excerpted randomly from various parts of the therapy sessions. Samples have most often been taken from the middle third to latter third of the
sessions because pilot work has demonstrated that these portions of the therapeutic encounters involve the least irrelevant, nontherapy related conversation.

Often the rater selects a sample by starting the session at a prescribed place, listening till a new person talks, and continuing for the time duration specified, but with provision that the sample contain at least one therapist and one client statement. These samples are then usually coded, listed in random order, rerecorded and then given to the raters for evaluation (Truax and Mitchell, 1971).

The data from several pilot studies of the Wisconsin study suggest that one to three samples per interview, two to five minutes in length, will give reliable representations of the session. Bocchini, Farwell, and Hart (1960) found that three to five minute segments were more reliable than shorter segments. Kiesler, Mathieu, and Klein (1964) compared patient process ratings of two, four, eight, and sixteen minutes and found no difference in range or reliability of ratings between the longer and the shorter segments. The Wisconsin studies utilized four minute segments extracted at random from the latter half of the interview.

Van der Veen (1967) argues that segment selection is critical and may have been partly responsible for his inability to replicate the results of previous investigators. Kiesler, Klein, and Mathieu (1965) and Karl and Abeles (1969) both provided evidence that the time segment sampled is a critical factor when sampling some variables. Bergin and Garfield (1971) found no difference with regard to outcome concerning which session was sampled. However, therapeutic conditions on
earlier tapes were found to be more variable.

However, other researchers have disagreed with these procedures. Kiesler, et al. (1967), Moose and Mackintosh (1970), Vesprani (1969), and Gurman (1973b) have questioned Rogers and his associates contention that the therapeutic conditions of empathy, warmth and genuineness are primarily attitudinal in nature and reflect underlying "attitudes" or "basic feelings" (Rogers, et al. 1967, p.11). They question the stability of the therapeutic conditions and disagree with Rogers conception of them as invariable therapeutic conditions existing throughout a session in a steady state.

Gurman (1973b) explored the stability of these therapeutic conditions for one patient of three high and three low functioning therapists. He found that while high and low facilitative therapists tended to function at high and low levels respectively, both groups were variable in their functioning between sessions and particularly within sessions. In addition he found that therapeutic conditions tended to reach peak within session levels in mid-late to late portions of the therapy hour.

If therapeutic condition ratings are not stable from patient to patient, it may indicate that the ratings measure qualities of the patient or the patient-therapist dyad, rather than qualities of the therapist. Beutler, et al. (1973) was concerned with this problem. After rating eight therapists initial therapy sessions with fifty-four patients they concluded that although accurate empathy may not be a stable quality of the therapist, as had been assumed, it does seem to be a dyadic or relationship variable.

Chinsky and Rappaport (1970) have raised three questions concerning
the aforementioned procedures of rating one's level of accurate empathy from audio tapes of psychotherapy sessions. They suggest that: (1) the raters may be responding to general counselor characteristics rather than his empathy; (2) reliability estimates may be associated with the number of therapists being rated; (3) reliability scores may be inflated by a lack of independent judgments.

Bozarth and Kraupt's (1972a) research considered each of these criticisms carefully. They trained experienced therapists to rate 1200 3-minute taped segments that were selected randomly from the final half hour of taped psychotherapy sessions. An analysis of their findings revealed that: (1) the reliability of the Accurate Empathy Scale is not necessarily a function of the number of therapists evaluated; (2) a single randomly selected segment per therapist may be as reliable as more than one segment per therapist; and (3) randomization seems to effectively safeguard against rater bias resulting from rater knowledge of the chronological sequence (Bozarth and Krauft, 1972).

Shapiro (1968) was interested in determining whether audio-taping of therapeutic interactions produces a reliable representation of the central therapeutic variables that are then related to outcome. Consequently he compared judgments of therapeutic conditions as rated through audio, video, and audio-video channels. He found, that although verbal and visual cues were found to be generally good predictors of the whole interaction, the raters were predisposed toward the verbal ratings. He concludes that audio-taping of sessions is a reliable abstraction of the therapeutic interaction itself.

Truax (1966) was concerned about the possibility that a rater's
assessment of a particular therapeutic condition was contaminated because the rater heard the patient's responses. In other words, he was concerned that the measures of therapist characteristics so widely employed might merely reflect good therapeutic interaction, and not the therapist's behavior as had been so widely concluded. Truax explored this possibility by taking 50 tape-recorded segments randomly selected from the course of treatment for five patients with five different therapists. By employing careful editing procedures he was able to obtain only the therapist statement portion for the same fifty samples (the patient's responses were completely edited out). Ratings were then made on both sets, those with and without the patient's replies. Analysis of the ratings indicated no significant differences between the measurements of accurate empathy and nonpossessive warmth in the edited and nonedited samples for the different therapy cases, for the different sessions from which samples were drawn, and for the different raters. It therefore appears that measurement of the therapist's levels of therapeutic conditions is in general, not contaminated by the patient's response to his remarks.

Although a number of criticisms have been levelled at the methodologies used in these studies, it appears that further research has clarified and resolved many of them. Those cited here and others mentioned earlier in this review (i.e., the poor inter-correlations found between therapeutic conditions in a few studies, etc.) are important methodological questions. However, it does appear that ratings of audio tapes of therapeutic sessions produce reliable measures of the level of therapist conditions present within any therapeutic interaction.
PREDICTING CLINICAL EFFECTIVENESS: A REVIEW OF THE LITERATURE ON THE PREDICTION OF OUTCOME EFFECTIVENESS FROM NON-PROCESS VARIABLES

This review has concentrated to this point on summarizing those studies relating certain therapist conditions to the therapeutic outcome. In other words only those studies that have predicted outcome based on certain therapist characteristics (empathy, warmth, and genuineness) have been reviewed. Since a number of studies have attempted to employ other variables or factors as predictors of outcome, they seem worthy of mention and consequently will be reviewed in this section.

Unfortunately, predictive psychology is still a rather crude and inconsistent science and since no studies have been published predicting success in therapeutic tutoring, as before, the literature from related areas of endeavor (i.e., counselor education, therapy, and student teaching) will be reviewed and hopefully will provide some useful insights into what causes one therapeutic tutor to be an effective agent of change while another is an ineffective agent. Particular attention will be given to the psychological measures others have employed to predict effectiveness apriori, since some psychometric measures will be used in this study to predict outcome effectiveness of therapeutic tutors.

Studies Attempting to Predict Teacher Effectiveness

Schluck (1970) found the MMPI not to be an effective predictor of observer ratings of teaching style. Holden (1971) employed a large battery of tests to evaluate changes in attitudes and interpersonal relations of pre-service teachers participating in an early childhood education program, as a consequence of their training intervention. He
found significant attitudinal changes for some of his sub-groupings as measured by the Minnesota Teacher Attitude Inventory and also the Authoritarian Attitude Towards Children Scale. Raina (1970) used the Torrence Tests of Creativity to predict teacher success, as measured by supervisors ratings and found insignificant correlations.

Solomon, et al. (1964) found that student's learning of facts was significantly related to teacher clarity and expressiveness and also to lecturing ability. Student's gains in comprehension were related to teacher energy and flamboyance and to moderate permissiveness scores. Warmth of teachers influenced the student's attitudes towards them, but was not related to quantitative learning.

Heil (1961) used a battery of tests in an attempt to relate various psychological and personality characteristics of the teachers to their student's progress. The teacher's scores as measured by the Manifold Interest Schedule (MIS) were significantly related to the progress of various categories of children at different levels of development. The Assessing Children's Feelings tests determined four categories of children: Conformers, Opposers, Waverers, and Strivers. The teachers were categorized into the Turbulent, the Self-Controlled and the Fearful groups based on their stanine scores on the MIS. Academic growth was assessed by the Stanford Achievement Test. Self-controlled teachers were found to be the most effective, with respect to academic growth, for all categories of children. Fearful teachers were found to be the least effective. When IQ differences and differences among the schools were partialled out, conforming children made slightly more progress under Turbulent Teachers. Achievement in reading and spelling was most
facilitated by Self-Controlled Teachers and progress in Math and Science was best with the Turbulent Teachers.

Hatfield (1961) found that self-concept, as measured by a differential Q-sort technique, was positively correlated to trained observers ratings of student teacher success. Williams, et.al. (1966) employed an extensive battery of tests to predict student teacher success and found only G.P.A. to be predictive. Openshaw (1967) found the Minnesota Teacher Attitude Inventory to lack predictive validity with regard to observer rated teacher success. Christensen (1960) used his own scales and the Affect-Need Scale to assess the relationships between pupil achievement, pupil affect-need, teacher warmth, and teacher permissiveness. The only significant relationship reported was between teacher's warmth and vocabulary and math achievement as measured by the Iowa Test of Basic Skills.

Studies Attempting to Predict Effectiveness as a Counselor

The data on predicting effectiveness in counseling is even less conclusive. Petty (1970) found that dogmatism scores were not predictive of counselor trainee success, but did seem to discriminate progress in a counselor training curriculum. After seven months involvement in the training program, trainee's dogmatism scores significantly declined. Stoffer (1968) and Foulds (1971) both found that the Rokeach dogmatism scale did not discriminate between facilitating and non-facilitating beginning counselors. However, Milliken and Peterson (1967) found that effective counselor trainees, as rated by supervisors, were significantly less dogmatic than their less effective peers.
Others have found different measures to be significantly related to supervisors ratings of counselor effectiveness. Russo, et. al. (1964) found the best counselors to score high on deference and order and low on abasement and aggression on the Edwards Personal Preference Schedule. Wasson (1965) found the Wisconsin Relationship Orientation Scale to significantly discriminate between the top 1/3 and bottom 1/3 of NDEA guidance enrollees rated on counselor effectiveness. Watton (1970) found graduate and undergraduate grades, the MMPI, the Miller Analogies, the Taylor Manifest Anxiety Scale, the Allport-Vernon Lindsey Scale of Values, and stated counselor attitudes and theoretical positions to not be predictive of counselor effectiveness.

Studies Attempting to Predict Therapeutic Effectiveness

A considerable quantity of research, much of it already reviewed here, has been concerned with determining the overall outcome of psychotherapy and the determinants of change within effective therapy. Prediction of effectiveness before or during therapy training or before the initiation of the actual therapeutic contacts has been rather limited and ineffective. Obviously, in any conceptualization of the therapeutic process, three main influences can be hypothesized. These are the client or patient, the therapist, and finally the interaction of the above two. This section will attempt to review a few of the studies not already covered that have employed characteristics of the therapist and/or the patient in predicting psychotherapeutic outcomes.

A. Client Variables Related to Therapeutic Outcomes

A variety of investigations have attempted to relate different client
attributes to selected outcome variables. Among those client variables that have been explored are: social class, personality variables, diagnostic categories, age, intelligence, motivation, expectation of change, attitudes, etc.

A number of studies have led to the conclusion that the dynamic, conventional long-term forms of psychotherapy are ineffective with disadvantaged patients of lower socioeconomic status (Hollinghead and Redlich, 1958; Imber, Nash, and Stone, 1955; Cole, Branch, and Allison, 1962; Rosenthal and Frank, 1958; and Rubenstein and Lorr, 1956) and that members of the lower socio-economic levels are less likely to be referred or to accept referral for psychotherapy (Rosenthal and Frank, 1958; Brill and Storrow, 1960; Schaeffer and Meyers, 1954; Cole, Branch and Allison, 1962). Clearly new procedures and innovative approaches are needed to deal with the plethora of mental health problems currently besetting our disadvantaged population.

Rosenberg (1954) examined the responses to the Wechsler-Bellvue, Rorschach and a sentence completion test for forty patients who received psychotherapy in an outpatient VA clinic. He found the successful patient to be of superior intelligence, flexible, sensitive, and active. However, Roberts (1954) and Rogers and Hammond (1953) were unable to find any correlations between their patient's Rorschach responses and their improvement rate.

Frank (1959), Rosenthal and Frank (1956), Frank, et al. (1959). Lipkin (1954) and others have found that the clients beliefs or expectations about the therapeutic process have a tremendous influence on their improvement rates. Friedman (1963) found a direct relationship
between expectancy and symptom reduction for his 43 patients. Goldstein (1960) found a positive relationship between expectations of improvement and their judged improvement by objective raters.

Client-therapist similarity has also been examined to determine its relationship, if any, with various outcome criteria and the findings are at best contradictory. Carson and Heine (1962) used the MMPI to compare therapists with their clients on varied psychological dimensions. They found a curvilinear relationship between therapist-client similarity and the client's rate of improvement. Lichtenstein (1966) and Carson and Llewellyn (1966) failed to replicate these findings, however. Lesser (1961) employed a Q-sort procedure and found that similarity in the client-therapist self-concept was negatively related to therapeutic process, while Levinson and Kitchener (1966) using a different Q-sort procedure attained more positive results. Luborsky, et.al. (1971) found, after reviewing fourteen articles, that only nine of the fourteen studies found a positive relationship between client-therapist similarity and outcome. Apparently the effects, if any, of client-therapist similarity will have to await more definitive research findings before any clear conclusions can be drawn.

Barcon (1953), Casner (1950), Fiske (1964), Rosenberg (1954), Rioch and Lubin (1959) and others have found that patients with higher intelligence scores perform better in psychotherapy. This is not surprising if one conceptualizes of psychotherapy as a learning process and therefore those who learn most easily will do better and will inevitable make the greatest progress.

Barcon (1953) found that the patient's ego-strength was positively related to the therapeutic outcome. However, Gallagher (1954), Getter
and Sundland (1962), Taulbee (1958), Klopfer, et al. (1951) and others have found that ego-strength measures were not predictive of outcome. Other researchers have attempted to use various anxiety scales, the MMPI, etc., to isolate characteristics of the client that are predictive of outcome. One might expect that predictions of outcome in psychotherapy would improve when psychological tests are used. However, this has not proven to be the case. Fulkerson and Barry (1961) and Luborsky, et al. (1971) all conclude, after reviewing the literature, that those client variables having the strongest relationship to outcome have been non-test variables; psychological health, motivation, social assets, acuteness of illness at onset, duration of illness, degree of precipitating stress, etc.

Clearly the problem of determining the personality attributes of clients that are related to their therapeutic outcome is a complex endeavor, which can not be readily accomplished through simplistic univariate small scale investigations. The variations within the samples, the therapists and their orientations, the outcome criterion measures, etc. have only limited replications and contributed to the seeming disparity in the findings published to date. It seems clear however that there are reliable and persistent changes in clients resulting from psychotherapy, but to date none of our measures of initial client status or initial client assets account for any very appreciable proportion of the variance of those changes.

B. Therapist Variables Other than Empathy, Warmth, and Genuineness Related to Therapeutic Outcome

Research efforts aimed at isolating those variables in the therapeutic process that effectively alter maladaptive behavior must inevi-
tably consider the therapist. He should be considered to be the primary agent of change. His skills and knowledge causes the client to change for better or worse. Some determinants of change have already been considered earlier and consequently will not be reviewed again here. However, other attempts at relating various therapist variables to outcome, will be reviewed here.

One approach that has received considerable attention is the A-B categorization of therapists. After reviewing the literature several conclusions can be reached. Type A therapists are better with schizophrenic patients, while type B therapists are better with neurotic patients. However, we have no clear idea of what the A-B categories actually measure, nor what it is about these categories that leads to client changes. Consequently the system has been of little real predictive or clinical value (Betz, 1963; Whitehorn and Betz, 1954; Stephens and Astrup, 1963; McNair, Callahan, and Lorr, 1962; and others).

Frayn (1968) used a battery of tests to predict the psychiatric residents who would be rated highly as therapists by supervisors. He found that those rated as highly proficient attained significantly higher scores on the Psychopathic Deviate Scale of the MMPI and significantly lower scores on the Compulsivity Scale of the Minnesota Hartfor Personality Assay. McNair, et al. (1962) and Whitehorn, et al. (1957) found vocational interest to be related to therapeutic success. Lawton (1965) found positive attitudes toward role-related concepts to be correlated with effective performance by non-professional mental health workers.
Luborsky, et al. (1971) in reviewing articles relating the therapist's level of experience to outcome, found that eight of the thirteen articles reviewed showed a significant positive relationship between the therapeutic experience and the improvement of the patient. The other five studies found no significant relationship. Further research is needed to clarify the impact of the therapist's level of experience on his therapeutic effectiveness. Young (1973) found the race of the therapist, the race of the patient, and the sex of the therapist to be unrelated to his outcome measures.

Other investigators have explored the role of therapist reinforcement in eliciting client changes. Krumboltz (1963) consciously gave verbal approval to vocational counselors whenever they verbally indicated their intention to seek information. He found that this increased the frequency of this behavior, when their rate was compared to the rate of a non-reinforced control group. Truax (1968) found that the therapists' differential reinforcement of client self-exploration was related to an increase in this behavior and also to greater therapeutic improvement on the outcome measures. Patients who received higher levels of therapist-offered differential reinforcement for self-exploration using momentary differential levels of therapist-offered accurate empathy, non-possessive warmth, and genuineness as reinforcers showed greater overall self-exploration three months after cessation of treatment and also greater overall improvement, than did those who received low or negative levels of differential reinforcement from their therapist. In addition, the therapists' use of these reinforcers was found not to be predictive of the mean level of these conditions offered.
during his therapy contacts. Thus, these findings suggest that moment-to-moment variation in levels of therapeutic conditions can significantly alter a patient's behavior and, depending on how they are used, the overall outcome of the psychotherapy process itself. Further research will have to determine the exact role that differential reinforcement plays in the therapeutic process.

Bergin and Solomon (1970) found that verbal intelligence scores and scores on the psychology subscale of the Graduate Record Exam (GRE) were negatively related to empathic ability. In addition they found that the P and D scales of the MMPI were negatively related with empathic ability while the Dominance and Change subscales of the Edwards Personal Preference Inventory (EPPI) were positively related to empathic ability. Veprani (1969) and Bergin and Jasper (1969) were able to replicate these findings for the MMPI subscales, but not for the subscales of the EPPI.

Allen (1967) concluded that no personality "types" or "profiles" are predictive of therapeutic effectiveness. However, he suggests that some higher order concepts like "openness" or the degree of self-awareness of the therapist might differentiate the effective from the ineffective therapist. He found that openness (as measured by Rorschach performance) was related to supervisor ratings of therapeutic effectiveness. The open counselors were seen as responding more often to client's feelings.

Unfortunately therapist characteristics have been unable to account for very much of the variance in the prediction of outcome in psychotherapy. Only a few of the studied characteristics have consistently been found to be positively related to outcome variables. Among those characteristics which seem most predictive are: therapist warmth,
empathy, and genuineness.

It has become clear from the extensive psychotherapy outcome literature that therapists who are accurately empathic, nonpossessively warm and genuine are indeed effective. Therapists high on these dimensions are able to produce the greatest changes in their clients. This study will explore the relationship between these three conditions and some other potentially predictive psychometric measures of certain therapist characteristics with the outcome measures of a therapeutic tutoring program with poorly achieving disadvantaged primary school age (grades 1-3) children.

From the articles reviewed in this section, it can be seen that few clearly predictive measures of effectiveness have arisen from areas seemingly related to therapeutic tutoring. Many good psychological measures were not found to be predictive, possibly due to invalid dependent variable measures and those measures which did correlate significantly with outcome often accounted for very little of the variance. One wonders about the value of generalizing from studies in these areas to prediction of success as a therapeutic tutor. The majority of the subjects in the reviewed articles were college students or adults and consequently differ from this study's anticipated population. In addition, two other concerns about these studies seem worthy of mention. First, the vast majority of them failed to confirm their significant findings in a cross-validation study, and second, very few bothered to develop any rationale of why a certain test was included within their predictive battery. Hypotheses were developed, it seemed, after the
data were analyzed and when significant predictive correlations were found, the results were often over interpreted. In addition, one index which is predictive in one study is often not predictive, with the same dependent variables, in another. Hopefully this study will begin to clarify some of these relationships and help to isolate those variables that are reliably able to predict effectiveness as a therapeutic tutor.
STATEMENT OF THE PROBLEM

The poor achievement of the disadvantaged child has been widely documented although the full etiology of this problem is not yet completely known. However, it has become clear that remedial procedures on the whole have been only marginally successful, especially over the long term with this population. This has led some theorists to suggest that either a critical period exists for certain forms of cognitive development, beyond which environmental manipulations will produce little or no change, or that some form of genetic inferiority, which is relatively fixed, is causing the poorer cognitive performances found in the disadvantaged population. Either way, remediation is seen as impractical.

However, other conclusions can be drawn from our apparent inability to permanently alter the poorer cognitive functioning of disadvantaged children. Our approaches to date may be totally inadequate and ineffective and in need of revision. This can only be determined through systematic evaluations and analyses of the factors that are associated with change. Some data look hopeful and one can remain optimistic that with further evaluation and research we will find those procedures and methods of intervention that are effective, economical, and practical.

This study, although concerned indirectly with the overall effectiveness of the therapeutic tutoring program as a remedial technique, is primarily interested in determining the characteristics of the effective tutor. More specifically, it will attempt to determine if certain
characteristics of the tutors are related to changes in the children's achievement levels (principally in reading) and in their behavior as rated by their parents, teachers, and peers. A number of tutor characteristics or traits will be explored (process ratings of their levels of empathy, warmth, and genuineness and their self concept, internality and dogmatism scores) and then related to the outcome measures to determine whether these variables are predictive of the tutor's effectiveness.

Despite the problems cited in the literature review, the determination of the characteristics of an effective tutor and/or tutoring relationship is an ambitious and worthwhile goal. Carkhuff (1969) has noted that few remedial intervention programs have systematically assessed their effects in terms of process variables related in previous research to a variety of indexes of constructive client outcome (Carkhuff, 1969, p. 4).

If one can ascertain the characteristics which define and constitute the effective tutor, then potentially one can use this information to predict success as a therapeutic tutor a priori. Obviously, this has great value for screening and hiring potential candidates for the tutoring program. Further, if certain therapeutic conditions or interpersonal skills are shown to be associated with a desirable outcome, then possibly, as Truax and Carkhuff (1967) have suggested, these conditions or characteristics can be taught or sharpened in successive generations of therapeutic tutors, with resultant improvements in their overall success rates.
SPECIFIC HYPOTHESES

The various hypotheses of this study will be presented in three different categories: 1) general hypothesis, 2) hypotheses relating to the prediction of outcome from condition (process) variables; and 3) hypotheses relating to the prediction of outcome from psychometric evaluations of the tutor. The outcome measures referred to in the hypotheses include the following change scores (post-pre tutoring evaluation scores): the Reading Comprehension, Word Analysis and Vocabulary subscales of the Iowa Test of Basic Skills, the Bender-Gestalt, a Parent-Teacher Behavior Rating Scale\(^1\) (See Appendix E). Positive changes indicate desired improvement for all scores except the Bender-Gestalt. In the case of the Bender, a significant negative correlation usually indicates a relationship with the change scores indicative of improvement.

GENERAL HYPOTHESIS

1. Some tutors will be more effective than others in eliciting desired changes. Thus, there will be significant variability among the tutors on the outcome change scores. Hypotheses dealing with the source of this differential effectiveness will be presented below.

\(^{1}\)The Parent-Teacher Rating Scale is an unpublished measure developed by Dr. Ellen Rie of the Ohio State University. A manuscript describing this scale has been submitted for publication to the Journal of Consulting and Clinical Psychology.
HYPOTHESES RELATING TO THE PREDICTION OF OUTCOME FROM CONDITION (PROCESS) VARIABLES

2. The combined ratings of tutor accurate empathic understanding on the Accurate Empathy Rating Scale (See Appendix A) will be significantly related to the children's improvement on the dependent variables.

3. The combined ratings of tutor warmth on the Nonpossessive Warmth Scale (See Appendix B) will be significantly related to the children's improvement on the dependent variables.

4. The combined ratings of tutor genuineness on the Genuineness Scale (See Appendix C) will be significantly related to the children's improvement on the dependent variables.

5. The sum of the combined rater summed Accurate Empathy, Nonpossessive Warmth, and Genuineness Scales will be significantly related to the children's improvement on the dependent variables.

HYPOTHESES RELATING TO THE PREDICTION OF OUTCOME FROM ASSESSED CHARACTERISTICS OF THE TUTOR

6. Tutors with a strong stable self-concept will be more effective therapeutic tutors in that they will produce greater academic and behavioral improvement in the children they tutor. Specifically, the Self Criticism Scores (SC), the Total P Score (TP), the P Score Identity (PI), the Total V Scores (TV), and the Personality Integration Scores (PI) for the tutors on the Tennessee Self Concept Scale will be predictive of the children's improvement on the dependent variables.

The SC score is a measure of openness and capacity for self-criticism. High scores indicate a normal healthy genuine openness. Extremely high scores indicate that the individual may be lacking in defenses and may in fact be pathologically undefended. Low scores
indicate inflexibility and defensiveness.

Persons with high scores on the TP scale like themselves, view themselves as helpful and valuable, and are confident of their abilities. Those with low scores have self doubts, see themselves as undesirable, often feel anxious, depressed and unhappy, and usually have little confidence in themselves.

The PI score reflects a person's basic opinion or view of himself. It is his description of his basic identity and consists of "what I am" items. The TV score provides a simple measure of a person's variability or inconsistency in his self perception from one area to another. A high score indicates that a person's self-concept is variable, reflecting little unity or integration. High scoring persons are often considered to compartmentalize certain areas of their self perceptions and to view these areas quite apart from the remainder of their self-concept. The better integrated person tends to score below the mean on this scale, but above the first percentile.

The PI score consists of items that discriminate individuals judged to have average or better levels of adjustment and personality integration. Higher scores indicate higher levels of adjustment or personality integration.

7. Post training scores for the tutors on the Rokeach Dogmatism Scale, Form E (See Appendix F) will be predictive of the children's improvement on the dependent variables. Higher scores on this scale indicate a higher level of dogmatism.

8. Tutor scores on Rotter's Internal-External Scale (I-E; see Appendix G) will be predictive of the children's improvement on the
dependent variables. Lower scores indicate a more internal locus of control.

9. Post training scores for the tutors on the Hogan Empathy Scale (see Appendix H) will be predictive of the children's improvement on the dependent measures. Higher scores are indicative of more empathic tutors.
CHAPTER II

METHOD

This study is an attempt to determine if certain characteristics differentiate the effective from the ineffective tutor.

This section describes the therapeutic tutoring program, the tutors, the children receiving tutoring, the measures, the assessment procedures utilized, and the experimental procedures for the process study.

THE THERAPEUTIC TUTORING PROGRAM

The reluctance of educators to concern themselves with the affective side of a child's development, the poorly achieving disadvantaged child's often disappointing response to traditional psychotherapeutic interventions, the increasing awareness of the cumulative deficit phenomena caused by severe inhibitions in achievement skills and an expanding interest in relating cognitive processes to affective components have led some to develop therapeutic tutoring programs (Rie, 1973; Templeton, Sperry and Prentice, 1967; and Trusty, 1971).

The therapeutic tutoring program with which this study is concerned was developed at the Children's Hospital, Columbus, Ohio, by Dr. Herbert Rie who directed the program (the author of this study worked as program coordinator). This program attempts to help the child improve both his achievement levels and his behavior by concentrating not only on his academic material, but also on the emotional component of his learning difficulties. In theory, the therapeutic tutor helps
the child more effectively deal with and understand his anxious
expectations and confused impressions of schools and teachers and through
the strength of the therapeutic tutoring relationship enables him to con­
front and disconfirm his learning inhibitions. As this is accomplished
the child is able to make better use of his academic opportunities and
his performance on his academic work will improve. This is not to imply a
short term process, but rather a continued period of growth and inter­
action occurring throughout the tutoring process (For a fuller descrip­
tion of the tutoring program see Rie, 1973). The tutoring lasted for
four months.

THE THERAPEUTIC TUTORS

Selection and Assignment of Tutors

This program served eight inner city schools in Columbus, Ohio.
Most of these are high priority schools and are considered to be among
those with the greatest academic delay. Although this is an ongoing
program, this study was conducted during the 1972-1973 academic year.

The principal of each of the eight schools were asked, if they
wished, to make referrals of potential tutors. The only stipulation was
that the tutor had completed high school. Every attempt was made to
recruit indigenous workers and several of the tutors actually tutored in
the schools in which their own children attended.

Tutors were interviewed and hired by the director of the program.
His decision to hire an individual was based on their interview behavior,
practical concerns (transportation, time available, etc.) and on their
responses to several psychometric measures (to be reviewed later in the
measures section). Naturally every attempt was made to hire the best
people, but the selection process itself was not seen as limiting
the variability to any great degree because only two applicants were
not hired.

Tutors were allowed to indicate the school where they wanted to
work, which usually was the closest school to their home, and whenever
possible that assignment was made. Children were assigned randomly to
the tutors based on the amount of time they were able to work, which in
a few cases was affected by the availability of children and overall
fund limitations. The tutors were paid $3.00 per hour. All tutors were
female. Only two tutors had previous tutoring experience with children
other than their own. Only tutors seeing four or more children were
included in the analyses of the therapeutic condition data (TCA), since
it was felt that this was the minimal number of children essential if
her effectiveness was to be reliably determined. This contingency re-
sulted in the loss of only one tutor from the process analysis.

Fourteen tutors were included in this study. One tutor resigned,
but was not included in the analyses and her children were reassigned
to another tutor. However, these children were not included in the
analyses of the study since they were seen by two different tutors. In
addition, a volunteer who tutored only two children was excluded from
the data analyses. Only eleven of the fourteen tutors could be included
in the analyses of the therapeutic condition data (TCA group). One
tutor was ill and consequently was unable to record her sessions.
Another tutor's tapes were totally inaudible and the final tutor worked
with only two children and therefore did not meet the requirement of
four considered essential for a reliable measure of her effectiveness to
be determined. These three were included in the analysis of the psychometric data, however.

Table 1 gives the breakdown by tutors of the grade levels for the tutored children from both the psychometric data (PA) and therapeutic condition analyses (TCA) groups. The group of children utilized for the psychometric analysis ($N = 109$) included 27 first graders (24.8%), 39 second graders (35.8%), and 43 third graders (39.4%). The numbers in each grade were more nearly equal for the therapeutic condition analysis group ($N = 81$). The TCA group contained 24 (29.6%) first graders, 30 (37.0%) second graders, and 27 (33.4%) third graders.

Table 2 contains tutor demographic information and intelligence scores. The mean age of the tutors was 38.86 years with a standard deviation of 11.52 and a range of 22-58 years. Their mean educational level was 12.64, with a standard deviation of 1.34 and a range of 11-16. Their mean IQ score (Peabody Picture Vocabulary Test) was 108.5 with a range of 85-138. The mean family income was $7,400 - $7,600 with a range of from less than $3,000 to over $12,000. Eight tutors were black, six were white.

After being selected for the program the tutors were trained and in January of the academic year began tutoring their children in their respective schools. They worked with the children individually outside their classrooms. They were expected to see each child two times per week for thirty to forty minutes. The mean number of tutoring sessions for the PA group was 28.09 (standard deviation 8.07). The mean overall number of minutes tutoring received by each child was 1,156.05 (standard deviation 326.91) with a mean session length of 41.37 minutes.
<table>
<thead>
<tr>
<th>TUTOR</th>
<th>FIRST GRADE PA—TCA</th>
<th>SECOND GRADE PA—TCA</th>
<th>THIRD GRADE PA—TCA</th>
<th>TOTALS PA—TCA</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2—2</td>
<td>6—6</td>
<td>3—3</td>
<td>11—11</td>
</tr>
<tr>
<td>2</td>
<td>3—3</td>
<td>4—4</td>
<td>2—1</td>
<td>9—8</td>
</tr>
<tr>
<td>3</td>
<td>0—0</td>
<td>0—0</td>
<td>8—0</td>
<td>8—0</td>
</tr>
<tr>
<td>4</td>
<td>0—0</td>
<td>5—0</td>
<td>1—0</td>
<td>6—0</td>
</tr>
<tr>
<td>5</td>
<td>3—2</td>
<td>1—1</td>
<td>2—2</td>
<td>6—5</td>
</tr>
<tr>
<td>6</td>
<td>3—3</td>
<td>3—3</td>
<td>0—0</td>
<td>6—6</td>
</tr>
<tr>
<td>7</td>
<td>2—1</td>
<td>5—5</td>
<td>3—3</td>
<td>10—9</td>
</tr>
<tr>
<td>8</td>
<td>4—4</td>
<td>0—0</td>
<td>1—1</td>
<td>5—5</td>
</tr>
<tr>
<td>9</td>
<td>1—1</td>
<td>4—3</td>
<td>0—0</td>
<td>5—4</td>
</tr>
<tr>
<td>10</td>
<td>2—1</td>
<td>1—1</td>
<td>4—4</td>
<td>7—6</td>
</tr>
<tr>
<td>11</td>
<td>4—4</td>
<td>6—6</td>
<td>5—5</td>
<td>15—15</td>
</tr>
<tr>
<td>12</td>
<td>1—1</td>
<td>2—1</td>
<td>5—5</td>
<td>8—7</td>
</tr>
<tr>
<td>13</td>
<td>2—2</td>
<td>0—0</td>
<td>9—3</td>
<td>11—5</td>
</tr>
<tr>
<td>14</td>
<td>0—0</td>
<td>2—0</td>
<td>0—0</td>
<td>2—0</td>
</tr>
</tbody>
</table>

TOTALS 27—24 39—30 43—27 109—81

# Tutors PA=14; TCA=11
TABLE 2

AGE, RACE, EDUCATIONAL LEVEL, MEAN FAMILY INCOME AND INTELLIGENCE SCORES FOR THE TUTORS

<table>
<thead>
<tr>
<th>TUTOR</th>
<th>AGE (in years)</th>
<th>RACE</th>
<th>HIGHEST GRADE COMPLETED</th>
<th>IQ&lt;sup&gt;a&lt;/sup&gt;</th>
<th>MEAN FAMILY INCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>53</td>
<td>B</td>
<td>12</td>
<td>126</td>
<td>8,100-10,000</td>
</tr>
<tr>
<td>2</td>
<td>31</td>
<td>W</td>
<td>12</td>
<td>85</td>
<td>6,100-8,000</td>
</tr>
<tr>
<td>3</td>
<td>44</td>
<td>B</td>
<td>12</td>
<td>105</td>
<td>Below 3,000</td>
</tr>
<tr>
<td>4</td>
<td>32</td>
<td>W</td>
<td>13</td>
<td>128</td>
<td>Below 3,000</td>
</tr>
<tr>
<td>5</td>
<td>37</td>
<td>B</td>
<td>13</td>
<td>129</td>
<td>4,600-6,000</td>
</tr>
<tr>
<td>6</td>
<td>49</td>
<td>W</td>
<td>11</td>
<td>115</td>
<td>10,100-12,000</td>
</tr>
<tr>
<td>7</td>
<td>22</td>
<td>B</td>
<td>12</td>
<td>101</td>
<td>Below 3,000</td>
</tr>
<tr>
<td>8</td>
<td>22</td>
<td>B</td>
<td>12</td>
<td>91</td>
<td>Below 3,000</td>
</tr>
<tr>
<td>9</td>
<td>29</td>
<td>W</td>
<td>16</td>
<td>102</td>
<td>Over 12,000</td>
</tr>
<tr>
<td>10</td>
<td>58</td>
<td>W</td>
<td>12</td>
<td>96</td>
<td>4,600-6,000</td>
</tr>
<tr>
<td>11</td>
<td>47</td>
<td>B</td>
<td>13</td>
<td>102</td>
<td>8,100-10,000</td>
</tr>
<tr>
<td>12</td>
<td>34</td>
<td>B</td>
<td>12</td>
<td>111</td>
<td>10,100-12,000</td>
</tr>
<tr>
<td>13</td>
<td>51</td>
<td>B</td>
<td>12</td>
<td>90</td>
<td>3,000-4,500</td>
</tr>
<tr>
<td>14</td>
<td>35</td>
<td>W</td>
<td>15</td>
<td>138</td>
<td>Over 12,000</td>
</tr>
</tbody>
</table>

Means: 38.86, 8B:6W, 12.64, 108.50, 7,400-7,600

ST. DEV.: 11.52, 1.34, 16.49

RANGE: 22-58, 11-16, 85-138

<sup>a</sup> Peabody Picture Vocabulary Test
Tutoring continued until early June when the final post-tutoring evaluations of the children were conducted (see measures section). The time tutored data represent the tutoring received between the pre and post tutoring evaluations.

TRAINING AND SUPERVISION

After being selected the tutors were given thirty hours of training by the program director. The tutors met once weekly for two hour didactic and experiential training. The first half of their training consisted of the presentation of the basic therapeutic principles. While the last half of the training consisted of the presentation of specific tutorial guidelines. The training sessions consisted primarily of lectures followed by periods of discussion. There was some video taping of interactions with children followed by discussion and feedback from the program director (for a more complete description of the training see Rie, 1974).

Following the training period the tutors began working with their respective children. During this period the program director met with the tutors to discuss their progress and/or problems. In addition, the program director and coordinator were always available to discuss any particular problems a tutor was encountering over the telephone.

THE CHILDREN

Principals reached agreement with the program director on the number of children that could be tutored in each school. They then requested that teachers refer children who were having difficulty in reading. These children were then evaluated and if they were six
months or more delayed on the Comprehension subtest of the Iowa Test of Basic Skills, they were accepted for tutoring. The only other stipulation set forth was that the child not be mentally retarded.

The children who were referred were evaluated (see measures section) and if they met the above requirements were randomly assigned to a tutor in their particular school. Children could be referred who were in other specialized programs, though one principal decided unilaterally against this policy.

In total, 117 children were tutored. However, only 109 of these were included in the PA group. Five children were tutored by two different tutors and consequently were omitted from the data analyses. Two other children were omitted from this group because their tutor was a volunteer who was not given the psychometric measures. One final child was deleted because he was a replacement for a child who moved and received a total of only 6.5 hours of tutoring (it was felt that this was a disproportionally small amount of tutoring). As a result, no child was included in the study who received less than ten hours of tutoring and/or ten tutoring contacts.

The TCA group consisted of 81 of the 109 children in the PA group. The 28 children who were omitted from this group were lost for various reasons. Fifteen could not be included because the tapes of their sessions were inaudible. Another 5 simply were not taped by their tutor because of their absences from school (the tutors had the recorders for only a limited period, partly because of a shortage of machines and partly to prevent the selective recording of sessions). Six more children were omitted because their tutor became seriously ill and was
unable to continue working. The final two subjects were deleted because their tutor worked with too few children.

Other variations in the samples will be found for each of the outcome variables. This is due to incomplete data for a particular child on a particular variable. For example, if a parent could not be contacted to complete the post tutoring PTQ this data would be omitted, though all the other data may be complete for that child. These variations in the sample will be noted as they occur in each group for the various outcome variables.

Table 3 contains demographic and intelligence data for the PA group categorized by their respective tutors. The mean age of the children in this group was 98.73 months, with a standard deviation of 12.02 and a range of 75-123 months (age at pre-tutoring evaluation). Their mean WISC Full Scale IQ was 89.46 with a range of 66-123. Sixty-one males and 48 females comprised this group, 60 were black, while 49 were white. Although specific social class data is not available for these children, they certainly were disadvantaged not only from the point of view of their academic skills, but also in their overall social status (income, parental level of education, welfare, etc.).

Table 4 consists of a tutor by tutor breakdown of the mean number of sessions tutored and the total mean hours spent tutoring each child in the PA group. Seventy-four of the children in the PA group were in no other specialized program. However, 35 members of this group received additional specialized inputs (Title I, speech therapy, Nisonger special reading teacher, psychological evaluation and consultation, etc.).
**TABLE 3**

CHILDREN'S DEMOGRAPHIC AND INTELLIGENCE DATA (PA GROUP) CATEGORIZED BY TUTORS

<table>
<thead>
<tr>
<th>TUTOR</th>
<th>MEAN AGE&lt;sup&gt;a&lt;/sup&gt;</th>
<th>SEX</th>
<th>RACE</th>
<th>MEAN IQ&lt;sup&gt;b&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>104.0</td>
<td>4M;7F</td>
<td>7B;4W</td>
<td>83.82</td>
</tr>
<tr>
<td>2</td>
<td>102.0</td>
<td>4M;5F</td>
<td>0B;9W</td>
<td>86.00</td>
</tr>
<tr>
<td>3</td>
<td>116.0</td>
<td>3M;5F</td>
<td>6B;2W</td>
<td>89.63</td>
</tr>
<tr>
<td>4</td>
<td>98.5</td>
<td>5M;1F</td>
<td>0B;6W</td>
<td>97.50</td>
</tr>
<tr>
<td>5</td>
<td>103.8</td>
<td>4M;2F</td>
<td>4B;2W</td>
<td>84.00</td>
</tr>
<tr>
<td>6</td>
<td>94.6</td>
<td>2M;4F</td>
<td>2B;4W</td>
<td>94.33</td>
</tr>
<tr>
<td>7</td>
<td>103.9</td>
<td>4M;6F</td>
<td>10B;0W</td>
<td>88.40</td>
</tr>
<tr>
<td>8</td>
<td>94.0</td>
<td>5M;0F</td>
<td>5B;0W</td>
<td>81.20</td>
</tr>
<tr>
<td>9</td>
<td>100.4</td>
<td>3M;2F</td>
<td>3B;2W</td>
<td>87.20</td>
</tr>
<tr>
<td>10</td>
<td>109.0</td>
<td>6M;1F</td>
<td>1B;6W</td>
<td>92.57</td>
</tr>
<tr>
<td>11</td>
<td>102.0</td>
<td>9M;6F</td>
<td>11B;4W</td>
<td>88.27</td>
</tr>
<tr>
<td>12</td>
<td>109.0</td>
<td>6M;2F</td>
<td>0B;8W</td>
<td>93.25</td>
</tr>
<tr>
<td>13</td>
<td>104.0</td>
<td>5M;6F</td>
<td>9B;2W</td>
<td>94.55</td>
</tr>
<tr>
<td>14</td>
<td>94.0</td>
<td>1M;1F</td>
<td>2B;0W</td>
<td>99.50</td>
</tr>
</tbody>
</table>

| MEANS  | 98.73 | 61'748F | 60B;49W | 89.46 |
| ST. DEVIAT. | 12.62 |         |         |       |
| RANGE   | 75-123|          | 66-123  |       |

<sup>a</sup> Age (in months) at pre-tutoring evaluation  
<sup>b</sup> Wechsler Intelligence Scale for Children (WISC) Full Scale score

N=109
### Table 4

The mean number of sessions and the mean number of hours of tutoring received per child

<table>
<thead>
<tr>
<th>Tutor</th>
<th>Mean Number of Tutoring Sessions per Child</th>
<th>Mean Number of Hours Tutoring Received per Child</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>31.27</td>
<td>20.40</td>
</tr>
<tr>
<td>2</td>
<td>33.20</td>
<td>23.75</td>
</tr>
<tr>
<td>3</td>
<td>29.38</td>
<td>20.97</td>
</tr>
<tr>
<td>4</td>
<td>24.33</td>
<td>18.71</td>
</tr>
<tr>
<td>5</td>
<td>20.50</td>
<td>13.67</td>
</tr>
<tr>
<td>6</td>
<td>22.67</td>
<td>16.83</td>
</tr>
<tr>
<td>7</td>
<td>21.60</td>
<td>14.13</td>
</tr>
<tr>
<td>8</td>
<td>51.60</td>
<td>31.05</td>
</tr>
<tr>
<td>9</td>
<td>26.00</td>
<td>13.95</td>
</tr>
<tr>
<td>10</td>
<td>26.43</td>
<td>20.21</td>
</tr>
<tr>
<td>11</td>
<td>29.13</td>
<td>19.02</td>
</tr>
<tr>
<td>12</td>
<td>25.75</td>
<td>18.97</td>
</tr>
<tr>
<td>13</td>
<td>26.45</td>
<td>21.05</td>
</tr>
<tr>
<td>14</td>
<td>28.00</td>
<td>20.88</td>
</tr>
</tbody>
</table>

**Means**  
28.09  
19.27  

**St. Dev.**  
8.07  
5.45  

*Number of tutoring sessions between pre and post tutoring evaluations*
After each prospective tutor was interviewed by the program director, she was asked to return at a later date at which time several psychometric measures were administered. Some of these measures were collected routinely as part of the project itself, while others were included as a portion of this study.

The psychometric evaluation of the tutors consisted of the following scales: Rotter's Internality-Externality (Locus of Control), the Tennessee Self Concept, the Rokeach Dogmatism, the Hogan Empathy and the Peabody Picture Vocabulary Test (these scales were administered to the tutors in the above order). Each of these scales will be reviewed below.

A. Rotter's Internal-External (Locus of Control) Scale (I-E)

This scale (See Appendix C) purports to measure a person's sense of being or not being in control of his life. When an event is perceived as resulting from luck, chance, fate or as unpredictable because of the great complexity of forces surrounding them, Rotter has labeled this as a belief in external control. However, if an event is viewed as contingent upon one's own behavior or characteristics this is considered by Rotter to be an internal belief system (Rotter, 1966, p.11).

A number of studies have found that individuals who score on the internal side are more likely to take social action to better their lives (Gore and Rotter, 1963), are more likely to attend to and learn information that will affect their future goals (Seeman, 1963), and are generally more concerned with their abilities, particularly when they fail (Elfram, 1964).
Harada and Scheibe (1967) found that internals placed greater value on skill or achievement reinforcements and scored higher on measures of dominance, defensiveness, endurance, and order. In addition, they found a positive correlation between I-E scores and a measure of maladjustment (a MMPI subscale) and also with a measure of anxiety (from the CPI).

Fitch (1970) found a positive correlation between a measure of self-esteem (Tennessee Self Concept Scale) and I-E scores. He also found a positive correlation between I-E scores and the Rokeach Dogmatism Scale. Joe (1971) found that externals were relatively anxious, aggressive, and dogmatic.

The scale is a forced choice questionnaire for which construct validity has been demonstrated in a variety of studies (Rotter, 1966; Joe, 1971; Seeman, 1963; Fitch, 1970, etc.) It has shown adequate internal consistency (.65 to .79) and test retest reliability (.49 to .83; Rotter and Mulry, 1965; Rotter, 1966). The scale samples general attitudes and the score consists simply of the number of external choices made by the respondent. The scale consists of twenty-nine items, including five fillers.

Concerning norms, Gore and Rotter (1963) found that 116 Black psychology students at Florida State University had a mean I-E score of 9.05 with a standard deviation of 3.66. Franklin (1963) found that a national stratified sample of 1,000 10th, 11th, and 12th graders achieved a mean I-E score of 8.50 with a standard deviation of 3.74.

B. Tennessee Self Concept Scale (TSCS)

This scale purports to measure the many facets of the self-concept.
It attempts to assess how an individual perceives himself and consequently has a wide variety of uses, in counseling, clinical work, research, etc. The scale consists of 100 forced choice self-descriptive items in which the respondent is asked to reveal his view of himself. The scale has two forms, a counseling and a clinical-research form. The latter was utilized in this study and is a more comprehensive form of the former. The various subscales of this test which were employed in this study were reviewed in Chapter I in the hypotheses section and will not be reconsidered here. The test is self-administered and requires a sixth grade reading level. Scoring instructions are available in the manual. Some scales (scores) are inverted.

This scale has demonstrated good test-retest reliability (.61 to .92; Fitts, 1965). Norms for the various subscales are presented in the manual (Fitts, 1965). Generally, respondent demographic variables have been found to be unrelated to the concept scores. The scale differentiates reliably between various diagnostic groups and consequently includes similarity scores (based on an item analysis) with these varying groups.

C. Rokeach Dogmatism Scale (Form E).

The Rokeach Dogmatism Scale (See Appendix F) attempts to measure differences in openness or closedness of belief systems. Rokeach (1960) defined the open belief system.

Every person, then must be able to evaluate 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 (Rokeach, 1960, p. 57).

The theoretical framework and a history of the development of the dogmatism scale is contained in Rokeach's book *The Open and Closed Mind* (1960). Form E of the scale contains forty items. Each item can be responded to on a six point scale ranging from strong agreement (+3) to strong disagreement (-3). The raw score for each item is computed by adding four points to the response value. Thus, a -3 becomes a +1, as negative scores are transformed into positive scores. The higher the total score, the more dogmatic or closed-minded the respondent.

Test-retest reliability is generally considered very adequate (.68 to .93; Rokeach, 1960). Concerning norms, Rabkin (1966) found that 107 American teachers in summer school achieved a mean score of 132.2 with a standard deviation of 22.5. There are no norms available for a group similar to the therapeutic tutors.

Finnigan (1971) found his open-minded subjects to be more self-reliant and less inclined to withdraw or manifest antisocial behavior. Kemp (1961) described highly dogmatic individuals as less aware of the feelings and desires at the center of their being. They tend, he suggested, to distort ideas that will not comfortably fit their particular belief system. If this is true, it would be difficult for a dogmatic tutor to be highly genuine. In addition, she would probably provide high levels of warmth only when the child shared her belief
system. Finally, it is very unlikely that unless one is able to understand his own belief system, that he would be able to understand, respect, and be empathic with the belief system of others.

Stoffer (1968) and Foulda (1971) found that dogmatism scores did not differentiate highly facilitative from ineffective non-facilitative beginning counselors. Parsons and Olsen (1969) found no relationship between the open-mindedness of the counselors and the levels of empathy they offered their clients.

Other studies however have reached different conclusions. Stefflre, et al. (1962) found that dogmatism scores differentiated effective from ineffective counselors. Milliken and Paterson (1967) found that good counseling trainees (as rated by their supervisors) were significantly less dogmatic than the poorer counselors. Saltzman (1966) observed that the degree to which a person is initially perceived by others to be warm, empathic, and genuine is a function of his level of dogmatism.

D. Hogan Empathy Scale

Hogan (1969) defines empathy behaviorally, as a willingness or tendency to modify one's behavior as the result of another's. He developed the scale because of the criticisms that had been forwarded concerning the other paper and pencil empathy measures. The item selection process is described by Hogan (1969). Primarily the items were selected based on the responses of high rated and low rated empathy groups to a pool of potential items. The scale itself (See Appendix H) consists of 64 self-descriptive items which are scored either true or false. The score simply represents the correspondence with the judged empathy group. The higher the score, the higher the level of empathy.
The scale consists of thirty-one items selected from the California Psychological Inventory, twenty-five from the Minnesota Multiphasic Personality Inventory (MPI) and the final eight items were derived from various experimental testing forms used at the Institute of Personality Assessment and Research in Berkeley, California.

The Empathy Scale, which was built to predict Q-sort derived empathy ratings, has good test-retest reliability (.68 to .86; Hogan, 1969). Concurrent validity scores with judged levels of empathy are also rather high (.62; Hogan, 1969). Hogan (1969) found the scale to be negatively correlated with the Rokeach Dogmatism Scale (-.31), positively correlated (.34) with ego strength (measured by the MMPI), and negatively correlated (-.40) with anxiety (measured by the MMPI).

Hogan found that 93 mixed level female college students mean score for the scale was 40.7 with a standard deviation of 5.9. Seventy female junior high school students (grade 7 and 8) achieved a mean score of 33.7 with a standard deviation of 5.2. Other norms more relevant to this study are not yet available.

E. Peabody Picture Vocabulary Test (PPVT)

The PPVT is a well known measure that was employed to assess the general intelligence levels of the tutors. The scale is an untimed individual test that is easily administered in fifteen minutes or less. It is purported to give an estimate of the subject's verbal intelligence and is highly correlated with other commonly used measures of intelligence (Strandberg, Griffiths, and Miner, 1969). Piers (1965) states that the PPVT is the best test of its kind now available.
The Dogmatism and Empathy Scales were readministered to the tutors after the completion of their training in an attempt to assess the impact of the training program. Petty (1971) found a decrease in counseling trainees' dogmatism scores after training, which suggested utilizing this scale as a measure of training progress. The initial overall mean empathy score ($N = 14$) was 37.50 (S.D. = 6.93). The follow-up post training empathy score was 35.43 (S.D. = 6.07), a change of -2.07. The change was more dramatic for the dogmatism scale. The initial mean score ($N = 14$) was 159.29 (S.D. = 29.90). The post training score was 149.14 (S.D. = 28.10), a change of -10.15. Thus, the training program lowered the dogmatism scores of the tutors considerably. This study will relate only the post-training empathy and dogmatism scores to the various outcome variables.

Tutor nonpossessive warmth, accurate empathy, and genuineness were assessed by utilizing the Truax scales (See Appendix A, B, and C) to rate taped segments of tutoring interactions (these procedures will be discussed in the next section). These scales were described previously and consequently will not be reconsidered here. Instead of the Truax (1961) Accurate Empathy Scale, the Bergin and Solomon (1970) revision was employed in this study. This modification includes a new stage between the old second and third stages.

Evaluation of the Children

The pre-tutoring assessment of those children referred to the program was conducted while the tutors were being trained during the last month prior to the actual start of the tutoring. The evaluation of the
children consisted of the following scales: The Reading, Vocabulary, and Word Analysis subscales of the Iowa Test of Basic Skills, the Wechsler Intelligence Scale for Children (WISC), the Bender-Gestalt, the Wesman Test of Auditory Discrimination, a behavior rating scale completed by their parent and teacher, and a classroom sociogram completed by the children's peers. Each of these scales will be reviewed briefly below.

A. Wechsler Intelligence Scale for Children (WISC)

The WISC is a widely used scale for measuring children's intellectual abilities. The data from this measure are complete for all subjects (N = 109). The entire test was administered initially with only the Arithmetic and Digit Span subscales being readministered during the post-tutoring evaluation at the end of the school year.

B. Iowa Test of Basic Skills (ITBS)

The ITBS reportedly measures children's basic achievement levels. Form five was used in this study and dependent on a child's grade, either level 7 or 8. The test was administered in small groups outside of the classroom. Two proctors administered the test. Three subtests of the overall scale were administered: Vocabulary, Word Analysis, and Reading Comprehension.

The Vocabulary subtest is a test of reading vocabulary. The child is required to read words that go with pictures and to comprehend the meanings of the word or concept. He is required to use decoding skills and to understand the meaning of words in a particular context.

The Word Analysis subscale measures a variety of skills involving sound letter association, phonetic analysis, and word structure. Stimuli consist of a variety of pictures, oral language, written language, and
The Reading Comprehension subscale includes three reading subtests. The first is a test of picture interpretation which contains pictures of explicit and implied actions and relationships. This task varies somewhat between level 7 and 8. The second subtest measures sentence comprehension and consists of yes and no answered questions. Vocabulary is carefully controlled to assure that the skill being measured is truly the understanding of relationships expressed in simple vocabulary. The final subtest is a test of story comprehension. It consists of several written passages followed by multiple-choice questions about each paragraph. The emphasis is upon understanding the ideas expressed or implied in the passage. The final reading comprehension score reflects performance on each of these subscales (Hieronymus and Lindquist, 1972).

Raw scores on each of these subscales were converted to grade equivalent scores for the purposes of this study. Graded percentile norms are presented in the manual. The data for these subscales is nearly complete. One child, because of absence, did not receive the final post-tutoring ITBS and consequently the number of complete ITBS was only 108. The post-tutoring assessment was done in the same manner as the pre-tutoring assessment.

C. Bender-Gestalt

This is a measure of visual (perceptual) motor performance. The Koppitz scoring system (a flaw score) was utilized for this study. In this test children were asked to replicate templates on another sheet of paper without tracing. The scale was administered individually when the WISC was given. It was also readministered during the post tutoring
evaluation. The complete N for the variable was 108.

Giebink and Birch (1970) and Connor (1967) found that Bender scores were generally ineffective predictors of reading achievement. However, Koppitz (1964) and Keogh (1969) both concluded after reviewing the literature that there was considerable evidence supporting the hypothesis that a relationship exists between Bender scores and academic achievement, especially in the primary grades where visual motor performance is so critical.

D. The Parent-Teacher Questionnaire (PTQ)

The PTQ (See Appendix D) is an unpublished behavior rating scale that is completed by parents (PTQ-P) and teachers (PTQ-T). It was recently developed by Dr. Ellen Rie of the Ohio State University Department of Pediatrics, Columbus, Ohio. It consists of thirty-five items which are rated in terms of their frequency of occurrence. A five indicates a very frequently occurring behavior, while a one indicates that the behavior never occurs. Although the scale is currently being factor analyzed, only the total score was utilized in this study. Other data are currently being collected on this scale and should soon be available. The higher the score the better the behavioral rating.

The scale was given to teachers to complete before tutoring began and again following the tutoring at the end of the academic year. The teacher was asked to either return the forms by mail (stamped envelopes were included) or to give them to the principal, from whom they were later picked up. Every attempt was made to get as many completed forms as possible (phone calls, letters, etc.). The parents completed the form twice, each time during a home visit by a member of the staff. The
first visit was to explain the nature of the program, to attain parental permission, and to collect this information. A follow-up visit, where the PTQ was readministered, was done following the cessation of tutoring (in June). The scale was read to each of the parents to avoid any literacy problems.

Inevitably, not all the data could be collected. The data for the teacher completed scale is complete (pre and post tutoring) for 89 children. It is somewhat better for the parent form where the data is complete for 101 children. Every attempt was made to have the same parent complete the form in each instance.

E. Classroom Sociogram (CS)

The CS (See Appendix E) is a measure of the child's popularity in his class among his peers. The form itself was completed by the teachers. The teachers were asked to have all the children in the class privately (on paper or verbally) indicate who they would like to sit next to. The children were asked to indicate first, second, and third choices. The teacher then simply indicated the number of times a tutored child had been chosen in each category on the form provided and returned it with the PTQ-T already discussed.

A total score was computed by summing three times the number of first choices, two times the number of second choices, and one times the number of third choices. Higher scores are indicative of a greater incidence of peer selection. Again the data is incomplete since 23 of the teachers did not return either the pre or post tutoring sociogram. (N = 86).
Glidwell and Swallow (1968) after reviewing several studies cite an average correlation of .47 between peers and teacher ratings of class members. Ullman (1952) found teacher's ratings of adjustment to be more closely related to sociometric ratings than to scores on self-descriptive tests. Many studies (Bower, 1960; Glidwell and Swallow, 1968; Cowen, Zax, Izzo, and Trost, 1967; and Zax, Cowen, Izzo, and Trost, 1964) have found that school achievement related significantly to the positiveness of ratings by teachers and peers. Bower (1960) states that the child most accepted by his peers gets the best grades, has the higher IQ, etc., etc.

F. Wepman Auditory Discrimination Scale (WADS)

The WADS (form 1) is a commonly used measure of auditory discrimination. The scale consists of forty items on which a subject is asked to determine whether two stimulus words are either the same or different. The X score is an error score reflecting the number of times the subject said same when the words presented were different. The Y score is simply the converse of the X score. The test was administered individually pre and post tutoring when the WISC and Bender scales were administered. Twenty-one pre or post tutoring scales were rejected as invalid because their scores were over the maximum number of errors allowed. This resulted in complete data on this variable for only 87 of the tutored children.

This study is concerned with predicting effectiveness as a therapeutic tutor from various psychometric characteristics of the tutors and from process ratings of the therapeutic sessions themselves. The
measures of effectiveness for the study are the change scores (post-tutoring minus pre-tutoring scores) for each of the outcome measures (ITBS, Bender-Gestalt, PTQ, and CS). Improvement is indicated by an increase in the change scores for the Iowa, PTQ, and the CS. However, a negative change score is indicative of improvement on the Bender-Gestalt and the Wepman scales.

There are many criticisms of utilizing change scores as measures of outcome. For example, one is forced to assume that changes occurring at different points reflect equivalent input (i.e., Is a grade equivalent change of from 2 to 2.7 equivalent to a change of from 2.7 to 3.4?). Despite the problems, change scores were selected as the outcome measures of change for this study.

EXPERIMENTAL PROCEDURE FOR THE PROCESS STUDY OF THE THERAPEUTIC CONDITIONS

Each tutor tutoring four or more children was required to tape one entire session with each of her children. Since they had recorders in their possession for only a short time, it was considered highly unlikely that selective taping would occur. The taping was done during the final third of the tutoring sessions, allowing for the stabilization of the relationship and for an equilibration of the number of sessions among the tutors. However, no attempt was made to rigorously assure that each child had received an identical number of tutoring sessions prior to being taped. Since the overall variability in number of sessions (See Table 4) was low and since the taping was done in the latter stages of the tutoring, this was deemed unnecessary. Bergin and Garfield (1971) concluded that in process studies no difference has
been found regarding outcome when different sessions were considered. However, therapeutic conditions in earlier sessions appear more variable, suggesting that studies should sample from the latter half of the therapeutic interaction.

The tutors were asked to assess on a four point scale (See Appendix I) the intersession representativeness of the audio-taped sessions (in comparison to other sessions with that particular child). In addition, she was asked when a session was seen as atypical, to describe as specifically as possible, how that session deviated from the norm. Table 5 contains the mean tutor ratings of the representativeness of the audio-taped sessions. After reviewing these ratings one becomes aware that the tutors saw the taped sessions as very similar (Mean 1.72, N=89) to their regular sessions with that child. In many cases recordings had already been used during prior tutoring sessions and the fact that this particular session was being taped was not considered to significantly affect the interaction. The directions given the tutor and the child for the taping are reviewed in Appendix I. Each tutor was instructed to follow her normal routine with her children during the session that was recorded.

Each tape was coded so that the raters could not identify either the tutor or the child. The entire tape was then reviewed so that academic rote interactions (i.e., reading) could be systematically eliminated from the selection process. Three segments were then selected from each tape, one from each third of the session (Gurman, 1973b had found considerable variability in therapeutic functioning within sessions, so it was decided that the segments should come from
### TABLE 5

**MEAN TUTOR RATINGS OF THE REPRESENTATIVENESS OF AUDIO-TAPED SESSIONS**

<table>
<thead>
<tr>
<th>TUTOR</th>
<th>FREQUENCY OF RELIABILITY RATING SCORES&lt;sup&gt;a&lt;/sup&gt;</th>
<th>N</th>
<th>MEANS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>1</td>
<td>11</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>3&lt;sup&gt;b&lt;/sup&gt;</td>
<td>2</td>
<td>4</td>
<td>2</td>
</tr>
<tr>
<td>4&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>5</td>
<td>4</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>4</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>7</td>
<td>1</td>
<td>5</td>
<td>2</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>11</td>
<td>9</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>14&lt;sup&gt;c&lt;/sup&gt;</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>MEAN OR TOTAL</strong></td>
<td>43</td>
<td>30</td>
<td>13</td>
</tr>
</tbody>
</table>

<sup>a</sup> 1= Very Similar; 2= Similar; 3= Different; 4= Very Different

<sup>b</sup> Tapes were inaudible

<sup>c</sup> No taping completed
different portions of the session). The first relevant segment contained in each third of the session was selected. In the few cases where there was no relevant segment available, one was selected from either the proceeding or following third of the tape. The segments were of two to four minutes duration and contained a minimum of one relevant tutor-child-tutor interaction. A few segments ran longer, to allow for closure and the inclusion of a complete interaction (Kiesler, et.al., 1964 found no difference in range, or sensitivity to individual differences when segment lengths of 2, 4, 8, and 16 minutes were employed in condition process studies).

The selection of segments was minimally limited by the number of relevant interactions available and to a much greater degree by inaudible or poorly recorded tapes. Twenty-eight children from the PA group were not included in the TCA group (N = 81). Fifteen of the 28 were because of inaudible tapes, while the other thirteen were never recorded for varying reasons. The number of tutors included in the TCA group was 11.

Two hundred forty three segments were selected from eighty one taped sessions. These segments were then numbered, placed in a randomized sequence, and rerecorded. Condition ratings of each segment were then made using Bergin and Garfield's (1970) ten point revision of the Truax Accurate Empathy Scale and Truax's five point scales of Nonpossessive Warmth and Genuineness. Each segment was then listened to by the raters and the three condition ratings were made in a counterbalanced order (Bozarth and Krauft, 1972 found that randomization effectively safeguarded against rater bias resulting from rater knowledge of the
specific chronological sequence). Two raters were employed in this study. The ratings were done separately. Each rater had a copy of the scales with a full description of each stage in front of him during the rating process. When variation existed within a segment, the modal condition score was recorded. If a rater had any doubts about how to rate a particular segment he could replay all or any portion of the segment before making his rating(s). All ratings were done blind with the rater not knowing either the tutor, the child, or the portion of the session sampled.

The raters were the author and his wife, who are both advanced graduate students at the Ohio State University in the clinical-child psychology program. Both raters had considerable clinical experience and had completed clinical internships. Both were in their twenties and were themselves eclectic in their therapeutic orientation.

The training procedure was identical for each of the raters and for each of the three rating scales. Each rater first studied the scales individually and became familiar with them and with the descriptions of each stage. The raters then discussed each of the scales in detail in an attempt to resolve any conceptual differences. Upon reaching agreement on the meaning of the scales and the stages the raters rated twenty practice segments. After hearing each segment, they discussed their scores and when differences appeared, examined them for possible theoretical biases. After twenty segments were rated, consistent agreement had been achieved and a common baseline established. The ratings of the actual segments were begun.
Table 6 presents the intercorrelations between the segments for each condition for each rater. Since these correlations were highly significant, indicating relative consistency (low variability) in the condition levels throughout the session, the three condition scores were summed for each rater and intercorrelated to determine the interrater reliability scores. These correlations (Pearson Product Moment Correlations) are presented in Table 7 along with each rater's mean condition scores. These correlations are high, indicating good interrater reliability, though well within the range found in other studies. These correlations ranged from a low of .86 for warmth to a high of .93 for the total summed condition interrater reliability scores. Additionally, the high correlations indicate that rater one's greater familiarity with the data (he reviewed the entire pool of tapes) did not bias his ratings in any discernable way.

Since the summed condition scores were intercorrelated (Table 7) they were combined (summed; rater 1 plus rater 2) and then related to the outcome variables to determine their predictive utility. Since the three conditions may operate jointly to influence the process movement and due to hypothesis five, a total overall condition (empathy, warmth and genuineness) score for the combined rater scores was also computed. The separate conditions approach however was considered to be a more stringent test of the hypotheses of this study.

Table 8 presents the intercorrelations of the conditions for the combined rater condition scores. These correlations are high and significant, though similar to the correlations found in other related process studies. However, their magnitude is sufficiently high to warrant some reconsideration of Roger's theoretical formulation of the
conditions as mutually independent traits of the therapist.
TABLE 6

INTERCORRELATIONS\* BETWEEN THE SEGMENTS FOR THE THERAPEUTIC CONDITIONS CATEGORIZED BY EACH RATER

<table>
<thead>
<tr>
<th>RATER ONE</th>
<th>ACCURATE EMPATHY</th>
<th>RATER TWO</th>
<th>ACCURATE EMPATHY</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SEG. 2</td>
<td>SEG. 3</td>
<td>SEG. 1</td>
</tr>
<tr>
<td>SEG. 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEG. 2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N=81

<table>
<thead>
<tr>
<th>RATER ONE</th>
<th>NONPOSSESSIVE WARMTH</th>
<th>RATER TWO</th>
<th>NONPOSSESSIVE WARMTH</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SEG. 2</td>
<td>SEG. 3</td>
<td>SEG. 1</td>
</tr>
<tr>
<td>SEG. 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEG. 2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N=81

<table>
<thead>
<tr>
<th>RATER ONE</th>
<th>GENUINENESS</th>
<th>RATER TWO</th>
<th>GENUINENESS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SEG. 2</td>
<td>SEG. 3</td>
<td>SEG. 1</td>
</tr>
<tr>
<td>SEG. 1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SEG. 2</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N=81

\* Pearson Product Moment Correlations
\*\*p < .01
\*\p < .001
<table>
<thead>
<tr>
<th>SUMMED CONDITIONS (THREE SEGMENTS)</th>
<th>RATER ONE (R1)</th>
<th>RATER TWO (R2)</th>
<th>INTER-RATER RELIABILITIES (R1 x R2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SUMMED ACCURATE EMPATHY (3 SECS.)</td>
<td>7.53 3.61</td>
<td>7.47 3.74</td>
<td>0.922*</td>
</tr>
<tr>
<td>SUMMED NONPOSSESSIVE WARMTH (3 SECS.)</td>
<td>7.94 2.92</td>
<td>7.80 2.88</td>
<td>0.868*</td>
</tr>
<tr>
<td>SUMMED GENUINENESS (3 SEGS.)</td>
<td>8.89 2.79</td>
<td>9.30 2.52</td>
<td>0.881*</td>
</tr>
<tr>
<td>SUMMED TOTAL CONDITIONS (9 SEGS.)</td>
<td>24.36 8.87</td>
<td>24.57 7.92</td>
<td>0.934*</td>
</tr>
</tbody>
</table>

N=81

a Pearson Product Moment Correlations

*p < .001
# TABLE 8

**INTERCORRELATION\(^a\) OF COMBINED RATER SUMMED CONDITION SCORES\(^b\)**

<table>
<thead>
<tr>
<th>COMBINED RATER SUMMED CONDITION SCORES</th>
<th>NONPOSSESSIVE WARMTH (R1+R2)</th>
<th>GENUINENESS (R1+R2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACCURATE EMPATHY (R1 + R2)</td>
<td>.8221*</td>
<td>.8187*</td>
</tr>
<tr>
<td>NONPOSSESSIVE WARMTH (R1 + R2)</td>
<td></td>
<td>.8845*</td>
</tr>
</tbody>
</table>

\(N = 81\)

\(^a\) Pearson Product Moment Correlations

\(^b\) These scores were derived from six rated segments, three from each rater.

\(*p < .001\)
CHAPTER III

RESULTS

This study was not concerned directly with the overall effectiveness of the therapeutic tutoring procedures and consequently it did not assess the program's overall success. Difference tests for within group (experimental) and for between group (experimental versus control) are not reported here. A manuscript containing this information is currently being prepared for publication at the Children's Hospital, Columbus, Ohio. The outcome change scores (post-pre tutoring evaluation) for each of the seven outcome variables were of considerable importance to this study, however.

OUTCOME CHANGE SCORES

Table 9 contains the mean pre-tutoring and mean change score data for the outcome variables for the PA group (N = 109). A positive change is indicative of improvement for each of the outcome variables, except the Bender-Gestalt, where a negative change score indicates that the child improved.

Table 10 presents the mean outcome change scores for each respective tutor. The second number in each cell represents the number of children from which the mean was derived.

Table 11 presents the correlations of the pre-tutoring scores for the outcome variables with the outcome change scores. With the exception of the Vocabulary Grade Equivalent score, the pre-scores...
## TABLE 9

**PRE-TUTORING MEANS AND MEAN CHANGE SCORES FOR THE OUTCOME VARIABLES**

<table>
<thead>
<tr>
<th></th>
<th>MEANS</th>
<th>STANDARD DEVIATION</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>PRE-TUTORING SCORES</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOWA VOCABULARY G.E.</td>
<td>1.212</td>
<td>0.667</td>
<td>109</td>
</tr>
<tr>
<td>IOWA WORD ANALYSIS G.E.</td>
<td>1.375</td>
<td>0.624</td>
<td>109</td>
</tr>
<tr>
<td>IOWA READING COMPREHENSION G.E.</td>
<td>1.066</td>
<td>0.866</td>
<td>109</td>
</tr>
<tr>
<td>BENDER-GESTALT</td>
<td>6.963</td>
<td>4.317</td>
<td>108</td>
</tr>
<tr>
<td>TOTAL PTQ-P</td>
<td>127.445</td>
<td>16.137</td>
<td>101</td>
</tr>
<tr>
<td>TOTAL PTQ-T</td>
<td>114.244</td>
<td>19.359</td>
<td>90</td>
</tr>
<tr>
<td>TOTAL CLASSROOM SOCIOGRAM</td>
<td>4.296</td>
<td>4.400</td>
<td>88</td>
</tr>
<tr>
<td><strong>OUTCOME CHANGE SCORES</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOWA VOCABULARY G.E.</td>
<td>0.562</td>
<td>0.567</td>
<td>108</td>
</tr>
<tr>
<td>IOWA WORD ANALYSIS G.E.</td>
<td>0.345</td>
<td>0.439</td>
<td>108</td>
</tr>
<tr>
<td>IOWA READING COMPREHENSION G.E.</td>
<td>0.791</td>
<td>0.646</td>
<td>108</td>
</tr>
<tr>
<td>BENDER-GESTALT</td>
<td>-1.500</td>
<td>3.196</td>
<td>108</td>
</tr>
<tr>
<td>TOTAL PTQ-P</td>
<td>5.347</td>
<td>12.717</td>
<td>101</td>
</tr>
<tr>
<td>TOTAL PTQ-T</td>
<td>7.322</td>
<td>15.353</td>
<td>90</td>
</tr>
<tr>
<td>TOTAL CLASSROOM SOCIOGRAM</td>
<td>1.302</td>
<td>5.204</td>
<td>86</td>
</tr>
</tbody>
</table>
**TABLE 10**

**MEAN OUTCOME CHANGE SCORES FOR EACH TUTOR**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.509</td>
<td>.418</td>
<td>.773</td>
<td>- .909</td>
<td>3.182</td>
<td>5.818</td>
<td>- .727</td>
</tr>
<tr>
<td>2</td>
<td>.544</td>
<td>.333</td>
<td>.767</td>
<td>- .889</td>
<td>2.889</td>
<td>1.778</td>
<td>.625</td>
</tr>
<tr>
<td>3</td>
<td>.763</td>
<td>.113</td>
<td>.338</td>
<td>- .250</td>
<td>8.125</td>
<td>20.875</td>
<td>- .500</td>
</tr>
<tr>
<td>4</td>
<td>.967</td>
<td>.250</td>
<td>.733</td>
<td>- 4.167</td>
<td>9.167</td>
<td>- 4.000</td>
<td>3.250</td>
</tr>
<tr>
<td>5</td>
<td>.383</td>
<td>.667</td>
<td>1.400</td>
<td>- 2.667</td>
<td>- 8.000</td>
<td>18.167</td>
<td>1.667</td>
</tr>
<tr>
<td>6</td>
<td>.250</td>
<td>.550</td>
<td>.333</td>
<td>.667</td>
<td>1.667</td>
<td>- 5.500</td>
<td>5.000</td>
</tr>
<tr>
<td>7</td>
<td>.722</td>
<td>.156</td>
<td>1.044</td>
<td>- 3.111</td>
<td>1.667</td>
<td>6.667</td>
<td>2.667</td>
</tr>
<tr>
<td>8</td>
<td>.460</td>
<td>.460</td>
<td>1.160</td>
<td>- 1.000</td>
<td>12.200</td>
<td>17.667</td>
<td>.667</td>
</tr>
<tr>
<td>9</td>
<td>.360</td>
<td>.320</td>
<td>.200</td>
<td>1.400</td>
<td>13.800</td>
<td>22.667</td>
<td>- 2.000</td>
</tr>
<tr>
<td>10</td>
<td>.557</td>
<td>.371</td>
<td>.886</td>
<td>- 3.286</td>
<td>2.857</td>
<td>- 1.000</td>
<td>.500</td>
</tr>
<tr>
<td>11</td>
<td>.313</td>
<td>.400</td>
<td>.853</td>
<td>- 1.800</td>
<td>10.000</td>
<td>6.500</td>
<td>.444</td>
</tr>
<tr>
<td>12</td>
<td>.850</td>
<td>.425</td>
<td>.825</td>
<td>.125</td>
<td>1.714</td>
<td>- 1.500</td>
<td>4.571</td>
</tr>
<tr>
<td>14</td>
<td>.800</td>
<td>.250</td>
<td>1.600</td>
<td>- 3.000</td>
<td>28.500</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>
### TABLE 11

**Correlations Between the Pre Tutoring Scores and the Outcome Change Scores for Each of the Outcome Variables**

<table>
<thead>
<tr>
<th>PRE TUTORING SCORES</th>
<th>OUTCOME CHANGE SCORES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IOWA VOC. G.E.</td>
</tr>
<tr>
<td>---------------------</td>
<td>-----------------</td>
</tr>
<tr>
<td>IOWA VOCABULARY G. E. PRE</td>
<td>0.036 (108)</td>
</tr>
<tr>
<td>IOWA WORD ANALYSIS G. E. PRE</td>
<td>0.271 (108)</td>
</tr>
<tr>
<td>IOWA READ. CO^p. G. E. PRE</td>
<td>0.221 (108)</td>
</tr>
<tr>
<td>BENDER-GESTALT PRE SCORE</td>
<td>-0.149 (107)</td>
</tr>
<tr>
<td>PTQ-P TOTAL PRE SCORE</td>
<td>0.163 (101)</td>
</tr>
<tr>
<td>PTQ-T TOTAL PRE SCORE</td>
<td>0.226 (90)</td>
</tr>
<tr>
<td>CLASS SOCIO TOTAL</td>
<td>0.092 (88)</td>
</tr>
</tbody>
</table>

* *p < .05; **p < .03; x*p < .01 (N in parentheses)

a Pearson Product Moment Correlations
were significantly and negatively correlated with their respective outcome change scores. Thus, as the pre-scores among the children increased, their outcome change scores tended to decrease. In other words, the children with the worst pre-scores tended to make the greatest improvement during the tutoring period. In addition, the Bender-Gestalt pre-score was significantly and positively related to the Iowa Reading Comprehension Grade Equivalent (GE) change score ($r = .204; p < .03, N = 108$). Thus, the children with the worst (highest) Bender-Gestalt pre-tutoring protocols tended to improve the most during the program on their Iowa Reading Comprehension GE scores.

No change scores were related to the time tutored variable contained in Table 4 (See Table 12). Consequently, the quantity and frequency of tutoring received was not controlled for, and the requirement that no child be included in the analyses who had not received a minimum of ten hours of tutoring in ten or more individual sessions apparently served as an adequate control for this potential source of variance.

Intercorrelations of the change scores revealed only the following significant relationships (See Table 13).

1. The Iowa Reading Comprehension GE change scores were negatively related to the Bender-Gestalt change scores ($- .279; N = 80, p < .022$). Thus, the children who improved the most in their Reading Comprehension GE scores also tended to improve the most on their Bender-Gestalt scores.

2. The Bender-Gestalt change scores were negatively related to the PTQ-T Total change scores ($- .243, N = 66, p < .025$). Thus, the teachers total ratings on the PTQ tended to show the greatest improve-
TABLE 12

CORRELATIONS BETWEEN THE MEAN NUMBER OF TUTORING SESSIONS RECEIVED PER CHILD AND THE TOTAL MEAN HOURS TUTORED PER CHILD WITH THE OUTCOME CHANGE SCORES

<table>
<thead>
<tr>
<th>CHANGE SCORES</th>
<th>N</th>
<th>MEAN NUMBER OF TUTORING SESSIONS PER CHILD</th>
<th>TOTAL MEAN HOURS TUTORED PER CHILD</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOWA VOCABULARY GRADE EQUIVALENT</td>
<td>108</td>
<td>.0525</td>
<td>.1272</td>
</tr>
<tr>
<td>IOWA WORD ANALYSIS GRADE EQUIVALENT</td>
<td>108</td>
<td>.0300</td>
<td>-.0284</td>
</tr>
<tr>
<td>IOWA READING COMP. GRADE EQUIVALENT</td>
<td>108</td>
<td>.0104</td>
<td>-.0136</td>
</tr>
<tr>
<td>BENDER-GESTALT</td>
<td>108</td>
<td>.0272</td>
<td>-.0108</td>
</tr>
<tr>
<td>PTQ-P TOTAL</td>
<td>101</td>
<td>.1628</td>
<td>.1344</td>
</tr>
<tr>
<td>PTQ-T TOTAL</td>
<td>90</td>
<td>.1291</td>
<td>.0071</td>
</tr>
<tr>
<td>TOTAL CLASSROOM SOCIOGRAM</td>
<td>86</td>
<td>-.0126</td>
<td>.0596</td>
</tr>
</tbody>
</table>

*Pearson Product Moment Correlations*
TABLE 13
INTERCORRELATIONS OF THE OUTCOME CHANGE SCORES a

<table>
<thead>
<tr>
<th></th>
<th>IOWA VOC. G.E.</th>
<th>IOWA W.A. G.E.</th>
<th>IOWA READ. COMP. G.E.</th>
<th>BENDER-GESTALT</th>
<th>PTQ-P TOTAL</th>
<th>PTQ-T TOTAL</th>
<th>CLASS. SOCIO. TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOWA VOC. G.E.</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOWA W.A. G.E.</td>
<td>0.097</td>
<td>(80)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOWA READ. COMP. G.E.</td>
<td>0.116</td>
<td>0.026</td>
<td>(80)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BENDER-GESTALT</td>
<td>-0.091</td>
<td>(80)</td>
<td>-0.070</td>
<td>-0.279**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTQ-P TOTAL</td>
<td>-0.236*</td>
<td>(73)</td>
<td>-0.034</td>
<td>0.072</td>
<td>0.072</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PTQ-T TOTAL</td>
<td>0.055</td>
<td>(66)</td>
<td>0.159</td>
<td>0.103</td>
<td>-0.243*</td>
<td>-0.150</td>
<td></td>
</tr>
<tr>
<td>CLASSROOM SOCIOGRAM-TOTAL</td>
<td>0.087</td>
<td>0.082</td>
<td>-0.13*</td>
<td>0.013</td>
<td>0.004</td>
<td>-0.084</td>
<td>1.00</td>
</tr>
</tbody>
</table>

*p < .03
**p < .01

a Pearson Product Moment Correlations
ment for the children who showed the greatest improvement on the Bender-Gestalt.

3. The Iowa Vocabulary GE change scores were negatively related to the PTQ-P Total change scores (-.236; N = 73, p<.022). Thus, the children who made the greatest improvement in their GE level on the Vocabulary subtest of the Iowa Test of Basic Skills tended to make the least improvement on the PTQ-P Total score.

ANALYSIS OF COVARIANCE FOR THE TUTORS WITH THE OUTCOME CHANGE SCORES WITH THEIR PRESCORES COVARED

Table 14 contains the analysis of covariance for the tutors with the outcome change scores with their pre-scores covaried. The pre-scores were held as covariates because of their significant negative relationship with the outcome change scores. With the effect of variation from pre-tutoring scores partialled out, no significant variation was found among the tutors for the Iowa Vocabulary and Word Analysis GE change scores and for the total classroom sociogram change scores (F = 1.046, 0.496, and 1.714 respectively, p >.05). The analysis of covariance for the total classroom sociogram change scores was nearly significant (p<.07). However, a significant variation was found between the tutors with pre-scores covaried for the Iowa Reading Comprehension GE change scores, the Bender-Gestalt, the PTQ-P Total change score, and the PTQ-T Total change score (F = 1.903, p<.04; F = 1.942, p<.04; F = 2.041, p<.04, and F = 2.824, p<.01). Thus, hypothesis one, that significant variation would be found between the tutors on the outcome change scores, can be accepted only for the above four variables.
### TABLE 14

**ANALYSES OF COVARIANCE**  
**TUTORS WITH THE OUTCOME CHANGE SCORES**  
**(WITH THEIR PRE SCORES COVARIED)**

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOWA VOCABULARY GRADE EQUIVALENT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error within</td>
<td>93</td>
<td>32.1853</td>
<td>1.046</td>
</tr>
<tr>
<td>Difference for testing adjusted means</td>
<td>13</td>
<td>33.7686</td>
<td></td>
</tr>
<tr>
<td>#Tutors = 14; #Cases = 108</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOWA WORD ANALYSIS GRADE EQUIVALENT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error within</td>
<td>93</td>
<td>18.6358</td>
<td>0.496</td>
</tr>
<tr>
<td>Difference for testing adjusted means</td>
<td>13</td>
<td>9.2487</td>
<td></td>
</tr>
<tr>
<td>#Tutors = 14; #Cases = 108</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IOWA READING COMPREHENSION GRADE EQUIVALENT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error within</td>
<td>93</td>
<td>27.4411</td>
<td>1.903*</td>
</tr>
<tr>
<td>Difference for testing adjusted means</td>
<td>13</td>
<td>52.2140</td>
<td></td>
</tr>
<tr>
<td>#Tutors = 14; #Cases = 108</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .04
<table>
<thead>
<tr>
<th>SOURCE</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>BENDER-GESTALT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error within</td>
<td>93</td>
<td>8.5743</td>
<td>1.942*</td>
</tr>
<tr>
<td>Difference for testing adjusted means</td>
<td>13</td>
<td>16.6477</td>
<td></td>
</tr>
<tr>
<td>#Tutors = 14; #Cases = 108</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL SCORE - PARENT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEACHER QUEST. - PARENT</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error within</td>
<td>86</td>
<td>122.1632</td>
<td>2.041*</td>
</tr>
<tr>
<td>Difference for testing adjusted means</td>
<td>13</td>
<td>249.3203</td>
<td></td>
</tr>
<tr>
<td>#Tutors = 14; #Cases = 101</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL SCORE - PARENT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TEACHER QUEST. - TEACHER</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error within</td>
<td>75</td>
<td>179.6100</td>
<td>2.824**</td>
</tr>
<tr>
<td>Difference for testing adjusted means</td>
<td>12</td>
<td>507.1340</td>
<td></td>
</tr>
<tr>
<td>#Tutors = 13; #Cases = 89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CLASSROOM SOCIOGRAM</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TOTAL SCORE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Error within</td>
<td>72</td>
<td>22.1265</td>
<td>1.714 (p &lt; .07)</td>
</tr>
<tr>
<td>Difference for testing adjusted means</td>
<td>12</td>
<td>37.9177</td>
<td></td>
</tr>
<tr>
<td>#Tutors = 13; #Cases = 86</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .04

**p < .01
Table 15 presents the mean combined rater summed condition scores for each of the tutors and also the overall means for the TCA group. The distribution of the condition ratings deserves some further discussion and analysis. To accomplish this, the condition scores for this study were compared to those reported by Rogers, et. al. (1967) and Stoffer (1968).

Figure 1 presents a comparison of the percentages that each stage of the Accurate Empathy Scale was rated in this therapeutic tutoring (TT) study, Stoffer's Community Helper (CH) program and Roger's Wisconsin Study. The latter two employed Truax's original scales and consequently no percentages are indicated for them for stage three. Each stage is described in Appendix A. Higher stages represent a higher condition level. Comparison of the percentages suggest a marked difference between the level of accurate empathy present in the therapeutic tutoring program and the Wisconsin study. The percentages are more similar for the TT and the CH programs. The distribution of percentage scores for the Wisconsin study was nearly normal, while it is highly skewed for the CH and TT programs.

Seventy-nine percent of the ratings from the TT program were assigned to the first (lowest) three stages. Sixty percent of the ratings for the CH program were assigned to the same three stages. However, only eighteen percent of the ratings from the Wisconsin study were assigned to these stages. This suggests that the therapeutic tutors when compared to the professional therapists in the Wisconsin study provided relatively low levels of rated empathy. However,
# TABLE 15

**TUTOR’S MEAN COMBINED RATER SUMMED CONDITION SCORES**

<table>
<thead>
<tr>
<th>TUTOR</th>
<th>NUMBER OF CHILDREN RATED</th>
<th>ACCURATE EMPATHY (R1&amp;R2)</th>
<th>NONPOSSESSIVE WARMTH (R1&amp;R2)</th>
<th>GENUINENESS (R1&amp;R2)</th>
<th>TOTAL CONDITIONS (R1&amp;R2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>13.818</td>
<td>17.364</td>
<td>19.909</td>
<td>51.091</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>9.625</td>
<td>11.750</td>
<td>12.000</td>
<td>33.375</td>
</tr>
<tr>
<td>5</td>
<td>5</td>
<td>16.200</td>
<td>19.600</td>
<td>23.400</td>
<td>59.200</td>
</tr>
<tr>
<td>6</td>
<td>6</td>
<td>28.500</td>
<td>23.167</td>
<td>26.333</td>
<td>78.000</td>
</tr>
<tr>
<td>7</td>
<td>9</td>
<td>16.111</td>
<td>19.444</td>
<td>21.333</td>
<td>56.889</td>
</tr>
<tr>
<td>8</td>
<td>5</td>
<td>13.000</td>
<td>11.400</td>
<td>17.600</td>
<td>42.000</td>
</tr>
<tr>
<td>9</td>
<td>4</td>
<td>17.750</td>
<td>11.750</td>
<td>17.250</td>
<td>46.750</td>
</tr>
<tr>
<td>11</td>
<td>15</td>
<td>8.600</td>
<td>9.400</td>
<td>12.200</td>
<td>30.200</td>
</tr>
<tr>
<td>12</td>
<td>7</td>
<td>22.857</td>
<td>22.000</td>
<td>24.143</td>
<td>69.000</td>
</tr>
<tr>
<td>13</td>
<td>5</td>
<td>18.000</td>
<td>20.400</td>
<td>18.800</td>
<td>57.200</td>
</tr>
</tbody>
</table>

| MEANS | 7.4 | 15.000 | 15.741 | 18.185 | 48.926 |
| ST. DEV. | 6.708 | 5.605 | 5.155 | 16.510 |

*# Tutors = 11; # Children = 81*

The above individual condition scores are the sum of the ratings of six segments, three derived from each rater. The total condition scores reflect the sum of the ratings of eighteen segments, nine from each rater.
PERCENTAGES OF THE RATINGS ASSIGNED TO EACH STAGE OF THE ACCURATE EMPATHY SCALE FOR THE THERAPEUTIC TUTORING PROGRAM, THE WISCONSIN STUDY, AND STOFFER'S COMMUNITY HELPER PROGRAM

The Community Helper Program and the Wisconsin Study used only the original nine point Truax scale and not the ten point revision by Bergin & Solomon employed in this study. The ten was not selected in any of the studies and consequently is omitted above.
although lower, their levels were similar to those provided by non-trained lay volunteers working with academically and behaviorally disturbed children (Grades 1-6) in a community helper program in Columbus, Ohio. In addition, some studies utilizing trained therapists have reported relatively low levels of rated accurate empathy (Bergin and Solomon, 1963; Melloh, 1964).

Figure 2 presents the percentages that each stage was used in rating Nonpossessive Warmth for the TT program, the CH program, and the Wisconsin Study. The stages are described in Appendix B. The higher stages indicate higher levels of rated nonpossessive warmth. Stage three was most often rated in each of the studies. However, only fifty-five percent of the ratings fell into the top three stages for the TT program, while ninety-three percent of the ratings from the Wisconsin Study were in this category (fifty-six percent for the CH program). Again, however, the nonprofessionals who were not trained in client-centered therapeutic techniques, in both the TT and CH programs, tended to provide lower condition levels than the trained therapists employed in the Wisconsin Study.

Figure 3 depicts the percentages that each stage was used in rating Genuineness for the TT program and the Wisconsin Study (The CH program did not report any genuineness scores because of their low inter-rater reliabilities for this condition). Each stage is discussed in Appendix C. Again, a higher score indicates a higher rated level of genuineness. As before, the therapeutic tutors tended to be rated lower than the professional client-centered therapists.

Not surprisingly, these comparisons tend to suggest that untrained
FIGURE 2

PERCENTAGES OF THE RATINGS ASSIGNED TO EACH STAGE OF THE NONPOSSESSIVE WARMTH SCALE FOR THE THERAPEUTIC TUTORING PROGRAM, THE WISCONSIN STUDY, AND STOFFER'S COMMUNITY HELPER PROGRAM

[Bar chart showing percentages for each stage of the warmth scale for the three programs.]

Therapeutic Tutoring

Community Helper Program

Wisconsin Study
FIGURE 3

PERCENTAGES OF THE RATINGS ASSIGNED TO EACH STAGE OF THE GENUINENESS SCALE FOR THE THERAPEUTIC TUTORING PROGRAM AND THE WISCONSIN STUDY

<table>
<thead>
<tr>
<th>Scale Stages</th>
<th>Therapeutic Tutoring</th>
<th>Wisconsin Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>35</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>62</td>
<td></td>
</tr>
</tbody>
</table>

a Stoffer did not report genuineness scores from his Community Helper Program
and minimally trained (TT) lay workers exhibit lower levels of the three therapeutic conditions than do psychotherapists trained in client centered techniques. The therapeutic tutors exhibited slightly lower levels of these conditions than the untrained volunteers in the CH program. However, the community helper volunteers were considered to be more middle class than the therapeutic tutors, were not paid, and met with only one child. The full implications of these lower scores will depend on the relationship found between the condition data and the outcome change scores.

Table 16 contains the correlations between the combined rater summed condition scores and the outcome change scores (Pearson Product Moment Correlations). The following significant relationships were found:

1. Nonpossessive warmth was negatively related to the PTQ-P Total change score ($r = -.205, p < .05$). Genuineness was also found to be negatively related to the PTQ-P Total change score ($r = -.194, p < .05$). Consequently, the children tutored by the tutors who were rated as the most warm and genuine tended to be seen as changing (improving) the least as a result of their involvement in the program, by their own parents. The data from the teachers (on the PTQ-T Total), although not significant, was in the same direction.

2. Each of the rated therapeutic conditions was significantly and positively related to the total classroom sociogram change scores (Accurate Empathy, $r = .319, p < .005$; Nonpossessive Warmth, $r = .286, p < .01$; Genuineness, $r = .287, p < .01$; and Total Conditions, $r = .318, p < .005$). Thus, the children of the tutors rated highest on the therapeutic conditions tended to change the most (became more popular).
### Table 16

**Correlation of Combined Rater Summed Condition Scores With the Outcome Change Scores**

<table>
<thead>
<tr>
<th>CHANGE SCORES</th>
<th>N</th>
<th>ACCURATE EMPATHY (R1+R2)</th>
<th>NON-POSSESSIVE WARMTH (R1+R2)</th>
<th>GENUINENESS (R1+R2)</th>
<th>TOTAL CONDITIONS (R1+R2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOWA VOC. G.E.</td>
<td>80</td>
<td>.048</td>
<td>.040</td>
<td>.089</td>
<td>.061</td>
</tr>
<tr>
<td>IOWA W.A. G.E.</td>
<td>80</td>
<td>.121</td>
<td>.066</td>
<td>.158</td>
<td>.121</td>
</tr>
<tr>
<td>IOWA READ. COMP. G.E.</td>
<td>80</td>
<td>-.081</td>
<td>.060</td>
<td>.020</td>
<td>-.007</td>
</tr>
<tr>
<td>BENDER-GESTALT</td>
<td>81</td>
<td>.110</td>
<td>.013</td>
<td>.024</td>
<td>.056</td>
</tr>
<tr>
<td>PTQ-P TOTAL</td>
<td>73</td>
<td>-.074</td>
<td>-.205*</td>
<td>-.194*</td>
<td>-.160</td>
</tr>
<tr>
<td>PTQ-T TOTAL</td>
<td>66</td>
<td>-.151</td>
<td>-.106</td>
<td>-.038</td>
<td>-.110</td>
</tr>
<tr>
<td>TOTAL CLASSROOM SOCIO.</td>
<td>65</td>
<td>.319**</td>
<td>.286**</td>
<td>.287**</td>
<td>.318**</td>
</tr>
</tbody>
</table>

*# Tutors = 11  N, TCA Group = 81

* a Pearson Product Moment Correlations

*p < .05

**p < .01

The above individual condition scores are the sum of ratings of six segments, three derived from each rater. The total condition scores reflect the sum of the ratings of eighteen segments, nine from each rater.
on the peer rated classroom sociogram.

These findings necessitate the rejection of hypotheses two through five for each of the outcome change scores, with the exception of the total classroom sociogram. Generally these findings suggest that the children of tutors rated high on each condition score and/or on the total condition score did not show significantly greater improvement than did their peers, who were tutored by less facilitative tutors. However, hypotheses two through five can be accepted in the case of the Classroom Sociogram Total change score.

TUTOR PSYCHOMETRIC DATA

Table 17 presents individual tutor scores and the group means for each of the psychometric measures (including the intelligence scores) which were hypothesized to be related to the tutoring outcome. These include: the Tennessee Self Concept Scale (TSCS); Self Criticism, Total P Score, P Score Identity, Total V Score and Personality Integration subscales), the post training Dogmatism, Rotter's I-E, and the post training Hogan Empathy Scale.

The tutor's mean scores on these measures were generally similar to the mean scores found for the standardization populations. However, the variability (standard deviation) of the tutor's scores was found to be somewhat lower in several cases than that found for the normative groups. This could reduce the scales potential predictive utility, in that as the variability in a population decreases, it becomes progressively more difficult to predict or discriminate differences.

The tutor's mean Total P Score (a measure of the overall level of self-esteem) was 349.21 with a standard deviation of 18.98. The mean
<table>
<thead>
<tr>
<th>TUTOR</th>
<th>TENNESSEE SELF CONCEPT SCALE</th>
<th>POST TRAINING DOGMATISM</th>
<th>POTTER'S I-E</th>
<th>POST TRAINING HOGAN EMPATHY</th>
<th>PFABODY IQ SCORE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SELF CRITICISM</td>
<td>TOTAL P SCORE</td>
<td>P SCORE</td>
<td>TOTAL V SCORE</td>
<td>PERSONALITY INTEGRATION</td>
</tr>
<tr>
<td>1</td>
<td>45</td>
<td>332</td>
<td>114</td>
<td>48</td>
<td>18</td>
</tr>
<tr>
<td>2</td>
<td>33</td>
<td>393</td>
<td>134</td>
<td>46</td>
<td>5</td>
</tr>
<tr>
<td>3</td>
<td>25</td>
<td>362</td>
<td>138</td>
<td>49</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>43</td>
<td>349</td>
<td>126</td>
<td>42</td>
<td>13</td>
</tr>
<tr>
<td>5</td>
<td>30</td>
<td>353</td>
<td>129</td>
<td>33</td>
<td>17</td>
</tr>
<tr>
<td>6</td>
<td>35</td>
<td>362</td>
<td>127</td>
<td>39</td>
<td>16</td>
</tr>
<tr>
<td>7</td>
<td>35</td>
<td>345</td>
<td>125</td>
<td>32</td>
<td>15</td>
</tr>
<tr>
<td>8</td>
<td>41</td>
<td>329</td>
<td>127</td>
<td>54</td>
<td>8</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>369</td>
<td>142</td>
<td>51</td>
<td>10</td>
</tr>
<tr>
<td>10</td>
<td>38</td>
<td>343</td>
<td>130</td>
<td>55</td>
<td>9</td>
</tr>
<tr>
<td>11</td>
<td>32</td>
<td>337</td>
<td>127</td>
<td>50</td>
<td>9</td>
</tr>
<tr>
<td>12</td>
<td>37</td>
<td>322</td>
<td>120</td>
<td>52</td>
<td>9</td>
</tr>
<tr>
<td>13</td>
<td>36</td>
<td>333</td>
<td>123</td>
<td>53</td>
<td>12</td>
</tr>
<tr>
<td>TUTOR</td>
<td>SELF CRITICISM</td>
<td>TOTAL P SCORE</td>
<td>P SCORE IDENTITY</td>
<td>TOTAL V SCORE</td>
<td>PERSONALITY INTEGRATION</td>
</tr>
<tr>
<td>-------</td>
<td>----------------</td>
<td>---------------</td>
<td>------------------</td>
<td>---------------</td>
<td>-------------------------</td>
</tr>
<tr>
<td>14</td>
<td>38</td>
<td>360</td>
<td>126</td>
<td>32</td>
<td>12</td>
</tr>
<tr>
<td>MEANS</td>
<td>35.214</td>
<td>349.214</td>
<td>127.714</td>
<td>45.429</td>
<td>11.286</td>
</tr>
<tr>
<td>RANGE</td>
<td>25-45</td>
<td>322-393</td>
<td>114-142</td>
<td>32-55</td>
<td>5-18</td>
</tr>
</tbody>
</table>
for the normative group was 345.57 (standard deviation 30.70). Thus, the tutoring group overall possessed a slightly higher level of self-esteem than the mean level of self-esteem for the standardization group of the TSCS (N = 626). The mean P Score Identity (one’s view of his basic self) for the tutors was 127.71 (standard deviation 7.02). The mean for the TSCS standardization group was nearly identical, 127.10 (standard deviation 9.96). The mean dogmatism score for the therapeutic tutors was 149.14 (standard deviation 28.10) while the mean score for 30 NDEA guidance and counseling institute attendents was 130.6 (Saltzman, 1966) and 132.2 (standard deviation 22.5) for 107 American teachers in summer school (Rabkin, 1966). Thus, this tutoring group was considerably more dogmatic than either of these normative populations. The remainder of these comparisons for each scale can be made by referring to Appendix K which contains the means and standard deviations for the standardization groups for most of the scales summarized in Table 17.

Only the following significant relationships were found when the psychometric measures listed in Table 17 were intercorrelated (Pearson Product Moment Correlations, N = 14). Intercorrelations among the sub-scales of the TSCS were not reported here however.

1. The post tutoring dogmatism scores were negatively related to the post training empathy scores (r = -.568, p < .017). Consequently, as the level of dogmatism among the tutors increased their empathy scores tended to decrease.

2. The post training dogmatism scores were negatively related to the tutor’s IQ scores (PPVT, r = -.640, p < .007). As the level of dogmatism among the tutors increased their level of intelligence tended
to decrease.

3. The post training dogmatism scores were negatively related to the Personality Integration scores on the TSCS \((r = -.541, p < .023)\). Thus, dogmatic tutors tended to receive lower personality integration (considered to be a measure of personality adjustment) scores and consequently could be considered to be less well adjusted.

4. The post training empathy scores (Hogan Scale) were found to be positively related to the Personality Integration subscales of the TSCS \((r = .644, p < .007)\). Thus, as empathy scores increased the tutor's personality integration scores also tended to increase. Therefore, empathic tutors tended to be better adjusted.

5. The tutor's IQ scores were found to be positively correlated with their TSCS Personality Integration subscale scores \((r = .619, p < .009)\). Thus, as the tutor's intelligence scores increased so did their personality integration scores.

6. The tutor's IQ scores were found to be negatively related to their TSCS Total V scores \((r = -.614, p < .01)\). The Total V score is considered to be an indication of the unity or integration of one's self concept (variability of perception). Thus, as the tutor's intelligence scores increased the variability in their perception of themselves tended to decrease. Therefore, brighter tutors were found to have a more stabilized self-image.

7. The tutor's I-E scores were found to be negatively related to the TSCS P Score Identity subscale \((r = -.637, p < .007)\). This subscale reportedly measures one's perception of his own basic identity. Thus, internal tutors tended to have a stronger more confident self-
image while external tutors tended to be more doubtful about their identity and less confident of themselves.

Table 18 presents the correlations between the various tutor psychometric scores and the mean (average for each tutor) tutor outcome change scores ($N = 14$). The following significant relationships were found.

1. The Self-Criticism subscale of the TSCS was negatively related to the PTQ-T Total change score ($r = -0.572, N = 13, p < .02$). Thus, the children tutored by tutors viewed on the TSCS as possessing a healthy, normal openness and capacity for self-criticism tended to be rated by their teachers as making the least improvement.

2. The Total P Score Identity on the TSCS was positively related to the PTQ-T Total change score ($r = 0.467, N = 13, p < .05$). Thus, the children of tutors possessing confident stable perceptions about their basic identity tended to be seen by their teachers as having made the greatest progress on the PTQ-T. This provides partial support for hypothesis 6.

3. The Total V Score on the TSCS was negatively related to the Iowa Reading Comprehension GE change score ($r = -0.468, N = 14, p < .05$). Thus, the children of tutors with a rather variable, poorly integrated, and compartmentalized self concept tended to make the least improvement in their grade equivalent scores on the Reading Comprehension subscale of the Iowa Test. Children tutored by tutors with a better integrated and less variable self image made the greatest progress in their grade equivalent scores on the Reading Comprehension subtest of the ITBS. This finding partially supports hypothesis 6.
**TABLE 18**

**CORRELATIONS** between the tutor psychometric data and the mean tutor outcome change scores

<table>
<thead>
<tr>
<th>PSYCHOMETRIC DATA</th>
<th>IOWA VOC. G.E.</th>
<th>IOWA W.A. G.E.</th>
<th>IOWA READ. COMP. G.E.</th>
<th>BENDER-GESTALT</th>
<th>PTQ-P TOTAL</th>
<th>PTQ-T TOTAL</th>
<th>TOTAL CLASSROOM SOCIOGRAM</th>
</tr>
</thead>
<tbody>
<tr>
<td>TENNESSEE (TSCS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELF CRITICISM</td>
<td>.302</td>
<td>.108</td>
<td>.375</td>
<td>-.436</td>
<td>.069</td>
<td>-.572</td>
<td>.323</td>
</tr>
<tr>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(13)</td>
<td>(13)</td>
</tr>
<tr>
<td>TOTAL P SCORE</td>
<td>-.148</td>
<td>-.122</td>
<td>-.228</td>
<td>.187</td>
<td>.118</td>
<td>.062</td>
<td>-.215</td>
</tr>
<tr>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(13)</td>
<td>(13)</td>
</tr>
<tr>
<td>P SCORE</td>
<td>-.204</td>
<td>-.195</td>
<td>-.383</td>
<td>.298</td>
<td>.175</td>
<td>.467**</td>
<td>-.435</td>
</tr>
<tr>
<td>IDENTITY</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(13)</td>
<td>(13)</td>
</tr>
<tr>
<td>TOTAL V SCORE</td>
<td>-.116</td>
<td>-.063</td>
<td>-.468**</td>
<td>.365</td>
<td>-.016</td>
<td>.122</td>
<td>-.415</td>
</tr>
<tr>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(13)</td>
<td>(13)</td>
</tr>
<tr>
<td>PERSONALITY INTEGRATION</td>
<td>-.168</td>
<td>.373</td>
<td>.220</td>
<td>-.231</td>
<td>-.291</td>
<td>-.176</td>
<td>.307</td>
</tr>
<tr>
<td>POST TRAINING DOGMATISM</td>
<td>-.201</td>
<td>.008</td>
<td>.080</td>
<td>.214</td>
<td>-.022</td>
<td>.281</td>
<td>.072</td>
</tr>
<tr>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(13)</td>
<td>(13)</td>
</tr>
<tr>
<td>ROTTER'S I-E SCORE</td>
<td>.248</td>
<td>.083</td>
<td>.618*</td>
<td>-.220</td>
<td>.230</td>
<td>-.112</td>
<td>.023</td>
</tr>
<tr>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(13)</td>
<td>(13)</td>
</tr>
<tr>
<td>POST TRAINING HOGAN EMPATHY SCORE</td>
<td>-.515**</td>
<td>.320</td>
<td>-.145</td>
<td>.072</td>
<td>-.062</td>
<td>-.079</td>
<td>-.152</td>
</tr>
<tr>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(14)</td>
<td>(13)</td>
<td>(13)</td>
</tr>
</tbody>
</table>
N Children = 109  N Tutors = 14

(Parentheses indicate number of tutors)

* Pearson Product Moment Correlations

**p < .02

**p < .05

<table>
<thead>
<tr>
<th>PSYCHOMETRIC DATA</th>
<th>CHANCE SCORES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>IOWA VOC. G.E.</td>
</tr>
<tr>
<td>PEABODY IQ</td>
<td>.238 (14)</td>
</tr>
</tbody>
</table>
4. The tutor's I-E scores were positively correlated with the Iowa Reading Comprehension GE change scores \( (r = .618, N = 14, p < .009) \). Thus, the children of external tutors tended to make the greatest improvement in grade equivalent scores on the Iowa Reading Comprehension test. This finding contraindicates hypothesis 8.

5. The post-tutoring Hogan Empathy scores were negatively related to the Iowa Vocabulary GE change scores \( (r = -.515, N = 14, p < .05) \). Thus, the children of tutors receiving lower empathy scores tended to make the greatest improvement on their GE scores on the Iowa Vocabulary subtest. This finding does not support hypothesis 9. However, since the correlations between the Hogan and the other Iowa change scores are highly variable, because of the low N, and because of other researchers findings, this relationship should be accepted only with the gravest reservation and only with further replication.

No other significant relationships were found to exist between the change scores and the tutor psychometric data. Some trends, though non-significant were found. These can be ascertained by an examination of Table 18. However, because of the low number of subjects, these findings should be considered only as suggestive, unless confirmed by subsequent findings.

These results, except as indicated in 2 and 3 above, do not provide support for hypotheses 6 through 9. These hypotheses are generally disconfirmed by the findings of this study.

ANALYSIS OF VARIANCE OF THE THERAPEUTIC CONDITION RATINGS

Table 19 reports the results of the analyses of variance by tutors for the therapeutic condition ratings. These analyses revealed
### TABLE 19

**UNIVARIATE ANALYSIS OF VARIANCE BETWEEN THE TUTORS AND THE THERAPEUTIC CONDITION RATINGS**

<table>
<thead>
<tr>
<th>SOURCE</th>
<th>df</th>
<th>MS</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMBINED ACCURATE EMPATHY (R1 + R2)</td>
<td>70</td>
<td>15.103</td>
<td>16.836*</td>
</tr>
<tr>
<td>between groups</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>within groups</td>
<td>10</td>
<td>254.274</td>
<td></td>
</tr>
<tr>
<td>COMBINED NONPOSSESSIVE WARMTH (R1 + R2)</td>
<td>70</td>
<td>9.055</td>
<td>20.757*</td>
</tr>
<tr>
<td>between groups</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>within groups</td>
<td>10</td>
<td>187.965</td>
<td></td>
</tr>
<tr>
<td>COMBINED GENUINENESS (R1 + R2)</td>
<td>70</td>
<td>4.240</td>
<td>43.149*</td>
</tr>
<tr>
<td>between groups</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>within groups</td>
<td>10</td>
<td>182.942</td>
<td></td>
</tr>
<tr>
<td>COMBINED TOTAL CONDITIONS (R1 + R2)</td>
<td>70</td>
<td>63.875</td>
<td>27.138*</td>
</tr>
<tr>
<td>between groups</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>within groups</td>
<td>10</td>
<td>1733.418</td>
<td></td>
</tr>
</tbody>
</table>

*p < .001
significant variation among the tutors for each of the condition scores (combined Accurate Empathy, F = 16.836, p < .001; combined Nonpossessive Warmth, F = 30.757, p < .001; combined Genuineness, F = 43.149, p < .001; and combined total conditions, F = 27.138, p < .001). Since significant variation was found between the tutors on the rated level of therapeutic conditions offered, an examination of some variables that might be related to or predictive of that variance seems warranted.

RELATIONSHIP OF TUTOR PSYCHOMETRIC SCORES TO RATED LEVEL OF THERAPEUTIC CONDITIONS

Table 20 presents the correlations of combined rater summed condition scores with the tutor's psychometric scores. The following were the significant correlations found.

1. Tutors scores on the Personality Integration subscale of the TSCS were positively correlated with the combined Nonpossessive Warmth score (r = .641, p < .02), the combined Genuineness score (r = .692, p < .01) and the combined total condition score (r = .605, p < .03). Thus, tutors found to have the greatest personality integration (similarity with a group of 75 judged to have average or better personality adjustment) tended to be rated higher overall on the warmth, genuineness, and total condition ratings.

2. Tutors scores on the Psychoses subscale of the TSCS were negatively correlated with the combined Nonpossessive Warmth scores (r = -.616, p < .03), the combined Genuineness scores (r = -.567, p < .04), and the combined total condition scores (r = -.572, p < .04). Thus, tutors found to have the greatest similarity (on items of the TSCS which reliably differentiated that diagnostic group from others) with
TABLE 20

INTERCORRELATION{a} TABLE OF THE COMBINED RATER SUMMED CONDITION SCORES AND THE TUTOR PSYCHOMETRIC SCORES

<table>
<thead>
<tr>
<th>PSYCHOMETRIC DATA</th>
<th>COMBINED RATER SUMMED ACCURATE EMPATHY (R1+R2)</th>
<th>COMBINED RATER SUMMED NON-POSSESSIVE WARMTH (R1+R2)</th>
<th>COMBINED RATER SUMMED GENUINENESS (R1+R2)</th>
<th>COMBINED RATER TOTAL SUMMED CONDITIONS (R1+R2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TENNESSEE (TSCS)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SELF CRITICISM</td>
<td>-.061</td>
<td>.171</td>
<td>.110</td>
<td>.069</td>
</tr>
<tr>
<td>TOTAL P SCORE</td>
<td>-.075</td>
<td>.231</td>
<td>-.261</td>
<td>-.191</td>
</tr>
<tr>
<td>P SCORE IDENTITY</td>
<td>-.136</td>
<td>-.504</td>
<td>-.381</td>
<td>-.346</td>
</tr>
<tr>
<td>TOTAL V SCORE</td>
<td>-.250</td>
<td>-.467</td>
<td>-.498</td>
<td>-.415</td>
</tr>
<tr>
<td>PERSONALITY INTEGRATION</td>
<td>.425</td>
<td>.641**</td>
<td>.692***</td>
<td>.605**</td>
</tr>
<tr>
<td>PSYCHOSES</td>
<td>-.465</td>
<td>-.616**</td>
<td>-.567*</td>
<td>-.572*</td>
</tr>
<tr>
<td>POST TRAINING DOGMATISM</td>
<td>-.079</td>
<td>-.218</td>
<td>-.154</td>
<td>-.154</td>
</tr>
<tr>
<td>ROTTER'S I-E</td>
<td>-.295</td>
<td>-.049</td>
<td>.058</td>
<td>-.114</td>
</tr>
<tr>
<td>POST TRAINING HOCAN EMPATHY SCORE</td>
<td>.229</td>
<td>.138</td>
<td>.186</td>
<td>.196</td>
</tr>
<tr>
<td>PEABODY IQ</td>
<td>.378</td>
<td>.499</td>
<td>.652**</td>
<td>.526*</td>
</tr>
</tbody>
</table>

#Tutors = 11  #Children = 81

{a} Pearson Product Moment Correlations

*p < .05
**p < .03
***p < .01
a psychotic group (N = 100) tended to have the lowest rated levels of Nonpossessive Warmth, Genuineness, and Total Conditions.

3. Tutor's scores on the PPVT were positively correlated with the combined Genuineness scores (r = .652, p < .02) and the combined total condition scores (r = .526, p < .05). Thus, intelligent tutors tended to have the highest rated levels of Genuineness and of Total Conditions. The correlations of IQ with other condition scores were positive, but these relationships were not significant.

It is difficult to interpret these findings, especially without further replication. They may represent an artifact of the rating process itself, the biases of the raters, or they may indeed be characteristics of truly facilitative therapists. Further research will hopefully clarify these results. It is interesting to note that psychometrically evaluated empathy (Hogan) was not significantly related (r = .229, p > .05) to process ratings of empathy levels. In addition, dogmatism and locus of control scores were found to be unrelated to the process ratings of therapeutic conditions.

**RELATIONSHIP OF TUTOR DEMOGRAPHIC VARIABLES WITH THE OUTCOME CHANGE SCORES**

Table 21 presents the correlations of tutor demographic variables with the outcome change scores. The following significant relationships were found.

1. The tutor's educational level (highest grade completed) was positively correlated with the PTQ-P Total change score (r = .626, p < .01) and with the PTQ-T Total change score (r = .491, p < .05). Thus, the children of the tutors who had received the most education,
TABLE 21

THE RELATIONSHIP a BETWEEN TUTOR
DEMOGRAPHIC VARIABLES AND THE OUTCOME CHANGE SCORES

<table>
<thead>
<tr>
<th>OUTCOME CHANGE SCORES</th>
<th>N</th>
<th>AGE OF TUTOR (YEARS)</th>
<th>RACE^b (B=1;W=2)</th>
<th>EDUCATIONAL LEVEL (Highest Completed)</th>
<th>TUTOR'S MEAN FAMILY INCOME</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOWA VOC. G.E.</td>
<td>14</td>
<td>-.240</td>
<td>.003</td>
<td>.011</td>
<td>-.324</td>
</tr>
<tr>
<td>IOWA W.A. G.E.</td>
<td>14</td>
<td>.120</td>
<td>-.030</td>
<td>-.136</td>
<td>.330</td>
</tr>
<tr>
<td>IOWA READING COMP. G.E.</td>
<td>14</td>
<td>-.251</td>
<td>-.172</td>
<td>.117</td>
<td>-.070</td>
</tr>
<tr>
<td>BENDER-GESTALT</td>
<td>14</td>
<td>-.018</td>
<td>-.022</td>
<td>.030</td>
<td>.485*</td>
</tr>
<tr>
<td>PTQ-P TOTAL</td>
<td>14</td>
<td>-.217</td>
<td>.350</td>
<td>.626**</td>
<td>.307</td>
</tr>
<tr>
<td>PTQ-T TOTAL</td>
<td>13</td>
<td>-.242</td>
<td>-.426</td>
<td>.491*</td>
<td>-.172</td>
</tr>
<tr>
<td>TOTAL CLASS. SOCI OGRAM</td>
<td>13</td>
<td>-.087</td>
<td>.048</td>
<td>-.534*</td>
<td>-.018</td>
</tr>
</tbody>
</table>

a Pearson Product Moment Correlations
^b Point Biserial Correlations

*p < .05
**p < .01
tended to be rated by both their parents and their teachers as making the greatest improvements on the PTQ.

2. The tutor's mean family income level was positively correlated with the Bender-Gestalt change score \( r = .485, p < .05 \). Thus, the children of the tutors who made the greatest income tended to make the least improvement on the Bender-Gestalt.

3. The tutor's educational level was negatively related to the total Classroom Sociogram change score \( r = -.534, p < .05 \). Thus, the children of tutors who had received the least education tended to make the greatest improvement on the peer rated Classroom Sociogram.

Generally these findings tend to support the conclusion that tutor demographic variables contributed very little to causing the changes noted on the outcome measures. Obviously they appear to account for very little of the variance of the outcome change scores, with the possible exception of tutor educational level with the PTQ-T Total change score.

THE RELATIONSHIP OF CHILD CHARACTERISTICS WITH THE OUTCOME CHANGE SCORES

Table 22 presents the correlations between certain child characteristics and the outcome change scores. The mean for the PA group on the WISC Verbal IQ was 88.47 \( (N = 109) \). For the WISC Performance IQ the mean was 92.41 \( (N = 109) \). The remaining PA group means for the variables found in Table 12 can be located in Table 3. The following significant relationships were found.

1. The children's age (in months) was negatively related to the Iowa Reading Comprehension GE change score \( r = -.276, p < .01 \). Thus,
<table>
<thead>
<tr>
<th>OUTCOME CHANGE SCORES</th>
<th>N</th>
<th>AGE (IN MONTHS)</th>
<th>SEX (1=M; 2=F)</th>
<th>RACE (1=B; 2=W)</th>
<th>GRADE</th>
<th>WISC VERBAL IQ</th>
<th>WISC PERFOR. IQ</th>
<th>WISC FULL SCALE IQ</th>
</tr>
</thead>
<tbody>
<tr>
<td>IOWA VOC.G.E.</td>
<td>108</td>
<td>.059</td>
<td>.043</td>
<td>-.021</td>
<td>.169</td>
<td>.241*</td>
<td>.185**</td>
<td>.242*</td>
</tr>
<tr>
<td>IOWA W.A. G.E.</td>
<td>108</td>
<td>-.137</td>
<td>.067</td>
<td>.097</td>
<td>-.207**</td>
<td>-.013</td>
<td>-.041</td>
<td>-.031</td>
</tr>
<tr>
<td>IOWA READ. COMP. G.E.</td>
<td>108</td>
<td>-.276*</td>
<td>.048</td>
<td>-.114</td>
<td>-.283*</td>
<td>.109</td>
<td>-.040</td>
<td>.046</td>
</tr>
<tr>
<td>BENDER-GESTALT</td>
<td>108</td>
<td>.184**</td>
<td>-.044</td>
<td>.073</td>
<td>.143</td>
<td>.043</td>
<td>.051</td>
<td>.054</td>
</tr>
<tr>
<td>PTQ-P TOTAL</td>
<td>101</td>
<td>.005</td>
<td>-.185***</td>
<td>-.004</td>
<td>.036</td>
<td>.010</td>
<td>.117</td>
<td>.125</td>
</tr>
<tr>
<td>PTQ-T TOTAL</td>
<td>90</td>
<td>.039</td>
<td>.064</td>
<td>-.328*</td>
<td>.043</td>
<td>-.112</td>
<td>-.231**</td>
<td>-.195***</td>
</tr>
<tr>
<td>TOTAL CLASSROOM SOCIO.</td>
<td>86</td>
<td>.113</td>
<td>-.062</td>
<td>.106</td>
<td>.066</td>
<td>-.080</td>
<td>-.057</td>
<td>-.076</td>
</tr>
</tbody>
</table>

*a* Pearson Product Moment Correlations

*b* Point Biserial Correlations

* *p* ≤ .01
** *p* ≤ .03
*** *p* ≤ .05
younger children tended to make the greatest improvement in their grade equivalent scores on the Iowa Reading Comprehension test.

2. The children's age (in months) was positively related to the Bender-Gestalt change scores (r = .184, p < .03). Thus, younger children tended to make the greatest improvement on the Bender-Gestalt.

3. The children's sex was negatively related to the PTQ-P Total change score (r = -.185, p < .05). Thus, males were rated by their parents as making greater improvement on the PTQ-P Total score than their female counterparts.

4. The children's race was negatively correlated with the PTQ-T Total change score (r = -.328, p < .001). Thus, teachers saw the black pupils as improving more than the whites on the PTQ-T Total score.

5. The children's grade level was negatively correlated with the Iowa Word Analysis GE change score (r = -.207, p < .03) and with the Reading Comprehension GE change score (r = -.283, p < .01). Thus, children in the lower grades tended to make the greatest progress (improvement) on grade equivalent scores on both the Iowa Word Analysis and Reading Comprehension subtests.

6. The Iowa Vocabulary GE change scores were positively correlated with the children's WISC Verbal IQ (r = .241, p < .01), WISC Performance IQ (r = .185, p < .03) and WISC Full Scale IQ (r = .242, p < .01). Thus, the more intelligent children tended to make the greatest grade equivalent improvement on the Iowa Vocabulary subtest.

7. The PTQ-T Total change scores were negatively correlated with WISC Performance IQ (r = -.231, p < .03) and with the WISC Full Scale IQ (r = -.195, p < .05). Thus, the least intelligent children tended
to be seen by their teachers as making the greatest improvement on the PTQ-T.

A SUMMARY OF THE FINDINGS

The following results constitute the major findings of this study.

1. There was significant variation among the tutors in their overall effectiveness on four of the seven outcome measures (Iowa Reading Comprehension GE change score, the Bender-Gestalt change score, the PTQ-P Total change score and the PTQ-T Total change score) and consequently an attempt to differentiate the effective from the ineffective tutor seemed justified, especially in relation to these four outcome measures.

2. Briefly trained therapeutic tutors tended to provide lower levels of Genuineness and Nonpossessive Warmth and especially low levels of Accurate Empathy when compared to professional psychotherapists employed in the Wisconsin Study. However, despite the overall low mean condition levels, there was still significant variation among the tutors in the level of these conditions present during their tutoring interactions.

3. Generally a tutor who provided a relatively higher level of Nonpossessive Warmth was also rated as providing a relatively higher level of Genuineness and Accurate Empathy and vice versa by both raters.

4. The levels of rated Accurate Empathy, Nonpossessive Warmth and Genuineness both singly and in combination were generally not significantly related to the outcome variables. However, each therapeutic condition (Accurate Empathy, Nonpossessive Warmth, Genuineness and the Combined Conditions) was significantly related to the total
Classroom Sociogram change scores. Thus, the children who received the highest levels of empathy, warmth, and genuineness tended to improve the most on the peer rated Classroom Sociogram. These findings led to the rejection of hypotheses two through five for each of the outcome measures with the exception of the Classroom Sociogram total change score.

5. The children tutored by tutors rated as the most warm and genuine tended to be seen by their parents as making the least improvement on the PTQ-P.

6. Tutor self concept scores were generally not related to the outcome measures in the direction indicative of improvement, except as indicated below.

   a. The children tutored by tutors with confident stable perceptions about their basic identity tended to be seen by their teachers as having made the greatest improvement on the PTQ-T. This provides partial support for hypothesis 6.

   b. The children tutored by tutors possessing better integrated and less variable self perceptions tended to make the greatest progress in grade equivalent scores on the Iowa Test of Reading Comprehension. This finding also provides partial support for hypothesis 6.

Thus, with the exceptions noted above, the findings disconfirmed hypothesis 6.

7. Dogmatism scores were unrelated to the outcome variables. Thus, dogmatic tutors were not found to be less effective. Hypothesis 7 was therefore not supported by the findings.

8. Internal tutors were not found to be more effective and thus
hypothesis 8 was not confirmed.

9. Empathic tutors (paper and pencil measure) were not found to be more effective. Hypothesis 9 was therefore not accepted.

10. The intelligence levels of the tutors were unrelated to any of the outcome measures.

11. The following significant relationships were found when the tutor psychometric measures were intercorrelated. Dogmatic tutors tended to be less empathic, less intelligent, and less well adjusted. The more intelligent tutors also tended to be better adjusted and had a more stabilized self perception. Internal tutors tended to possess a more confident self-image, while empathic tutors tended to be better adjusted.

12. Tutors rated highest on Genuineness, Nonpossessive Warmth, and Total Conditions tended to have the highest levels of personality integration (adjustment).

13. The intelligent tutors tended to receive the highest levels of rated Genuineness and Total Conditions.

14. The children with the poorest pretutoring scores tended to make the greatest improvement on the outcome measures.

15. The children tutored by the tutors with the most education tended to be rated by both their parents and their teachers as making the greatest improvement on the PTQ.

16. The younger children tended to make the greatest improvement on the Iowa Reading Comprehension subtest and the Bender-Gestalt. Children in the lower grades tended to make the greatest improvement on the Iowa Word Analysis and Reading Comprehension subtests.
17. Males tended to be rated by their parents as making greater progress on the PTQ-P. The teachers rated the black children as making greater progress than whites on the PTQ-T.

18. Teachers tended to rate the least intelligent children as making the greatest improvement on the PTQ-T.
CHAPTER IV
DISCUSSION

Chapter four contains a discussion of the major findings of this study, some criticisms of the study, and suggestions for further research.

This study attempted to determine the characteristics that influence the outcome of a therapeutic tutoring program with poorly achieving inner city disadvantaged primary school age children (1st to 3rd grade). Because of the research findings which suggested that characteristics of the therapist are the primary determinants of change in any therapeutic relationship and because a therapeutic orientation is a part of this tutorial program, the tutor's level of Accurate Empathy, Nonpossessive Warmth, and Genuineness was assessed. In addition, the tutor's self-concept, locus of control, level of openmindedness (dogmatism), and intelligence were also measured and then related to the studies' outcome measures, in an attempt to differentiate the effective from the ineffective tutor. Demographic characteristics of the tutors and some child variables (age, sex, intelligence, etc.) were also related to the outcome scores in a further attempt to achieve a fuller understanding of the factors that are associated with (predictive of) changes in the children's academic and behavioral performance.

The measures of change (outcome) employed in this study consisted of a test of reading achievement (Iowa Test of Basic Skills Reading Comprehension, Vocabulary, and Word Analysis subtests), a measure
of perceptual motor ability (the Bender-Gestalt), a parent and teacher completed behavioral questionnaire, and a peer completed classroom sociogram. These scales were administered both prior to and following the tutoring period.

The specific purpose of this study was to investigate the hypotheses that high levels of rated Accurate Empathy, Nonpossessive Warmth and Genuineness in tutors, both singly and in combination would be significantly related to the outcome; that the tutor's self-concept would be significantly related to the outcome; that dogmatic tutors would be less effective; that internal tutors would be more effective; and that empathic tutors (derived from a psychometric measure rather than from process ratings) would be more effective.

DISCUSSION OF THE FINDINGS

It is difficult to account for the finding that generally the level of therapeutic conditions offered by the tutors tended to be unrelated to the outcome measures (with the exception of the Classroom Sociogram change score). Numerous investigations (reviewed in Chapter I) have found significant relationships between the levels of rated Nonpossessive Warmth, Accurate Empathy, and Genuineness and a variety of outcome measures for diverse patient populations receiving different forms of therapy. However, Garfield and Bergin (1971) found no relationship between the rated condition scores and the outcome measures when they studied process recordings from therapists who were not specifically trained in client-centered techniques.

Few studies have assessed the impact of therapeutic conditions on young children, and especially their effect on their cognitive skills.
It is not yet clear that young children's academic performance necessarily improves with counseling or psychotherapy, though their behavior certainly can be affected by a therapeutic relationship. Calhoun (1956) found no significant improvement for children on achievement tests resulting from their placement in a counseling group. In addition, those studies where children have been found to improve after receiving counseling or psychotherapy have often studied older (adolescent) and brighter populations, rather than poorly achieving young disadvantaged children. In fact, many theorists and researchers have concluded that the traditional therapeutic techniques are ineffective with disadvantaged children. Further research will hopefully clarify the value of a therapeutic relationship with a young disadvantaged child. The failure of facilitative therapists to produce greater changes in this population's academic performance is not surprising, since there is little basis for expecting an impact of therapeutic conditions on intellectual functioning in a deprived population. The impact of the therapeutic conditions with this group however certainly warranted study and exploration.

Van der Veen (1967) was unable to replicate some of the previous therapeutic process research findings. He felt that this resulted from variance introduced from the segment selection procedures. Most studies have selected tapes from the latter portions of the therapeutic interaction. Stoffer (1968) found that rated therapeutic conditions derived from tapes of early sessions were unrelated to his outcome measures, while rated conditions from tapes of later sessions were related to the outcome data. Bergin and Garfield (1971), Rogers, et al. (1967), and
others have found variability in the level of rated conditions during sessions. However, there are no real guidelines for therapeutic tutoring and it is possible that the session sampled may prove to be a crucial factor (the taping in this study was done near the end of the tutoring).

Gurman (1973b) found considerable within session variability in rated levels of therapeutic conditions, suggesting that a study's segment sampling procedure might be critical. The peak levels for each of the conditions, he found, tended to be reached in the mid-late to late portions of the session. He also found greater variability among inexperienced or less experienced therapists. Due to Gurman's findings and despite Bozarth and Krafft's (1972) conclusion that one randomly selected segment was as representative of the entire session as were several segments, one segment from each third of the session was selected. If Gurman's findings are indeed valid, perhaps the segments sampled in this study should have been selected from the latter half of the session, when intra-session variability was minimized and peak condition levels achieved. Since inter-segment correlations (See Table 6) were lower for Accurate Empathy, this condition score may be the most vulnerable (greatest variability). Truax (1962) found that the highest moments of Accurate Empathy obtained throughout a session were more predictive of the outcome than were the mean empathy levels or the levels from sessions with relatively lower highest moments. In addition, the low moments of empathy were not found to be predictive of outcome. Further research may clarify whether the inter-session segment selection procedure is crucial and whether a procedural modification in this study would have altered its findings. The segment selection procedure was
further complicated in this study by non-therapeutic rote academic interactions. Perhaps, in future studies of this type these interactions themselves should be quantified (percentage of the session spent in this type of activity) and studied carefully themselves (i.e., percentage of child successes, reinforcement schedule, etc.)

When intra-segment variability was observed (primarily when rating accurate empathy) the modal condition score was employed rather than the peak level. In future studies this procedure could be changed, especially considering Truax's (1962) and Gurman's (1973b) findings.

Perhaps the crucial element in therapeutic tutoring with children is condition specificity rather than the overall mean level of facilitativeness. The mean levels of conditions may be unrelated to the outcome, while condition specificity, where the therapist responds facilitatively at certain points in the therapeutic interaction that are critical (i.e., when a child is dealing with material most related to his symptom) may be highly related to the outcome. Therefore, these interactions may be much more important than the overall mean level of facilitativeness that was considered in this study.

The high intercorrelations between the various conditions (See Table 8) could be interpreted in several ways. Perhaps tutors who provide high levels of one condition are also likely to provide high levels of the other conditions. This would suggest that the levels of rated conditions reflect an underlying attitude of the tutor. However, these high correlations might also indicate that the various conditions are not independent traits, but rather either unidimensional characteristics of the tutors or a reflection of the biases of the raters.
Since the same raters rated each condition (even though the rating order was varied) they may have been influenced by their early ratings and thereby unknowingly their inter-rater reliability scores could have been inflated. However, these reliability scores were well within the range reported in other studies, though somewhat higher than the average. Consequently, it is difficult to assess the meaning of these correlations and thus nearly impossible to reflect upon Rogers' (1967) contention that these conditions reflect basic invariable attitudes of the therapist.

Zimmer and Anderson (1968) also questioned the assumption that the facilitative conditions reflect unidimensional orthogonally related personality traits of the therapist. After factor analyzing 100 counselor responses, 8 principle factors were identified for Nonpossessive Warmth and Accurate Empathy. Consequently, Traux's procedure of viewing these conditions on a continuum from more to less may be totally invalid. Further factor analyses of therapeutic condition data may clarify this possibility and produce a complete change in our conceptions of the therapeutic conditions themselves.

Overall, therapeutic tutors mean condition scores were rather low when compared to the levels of professional therapists, especially on Accurate Empathy. It seems possible that even though there was significant variation among the tutors on the levels of conditions they offered the children, the levels were low enough to preclude constructive changes in achievement levels, but high enough to have some impact on the behavioral measures. Perhaps, as Gendlin and Geist (1962) suggest
since one condition score was generally lower than the other two, there might have been a concomittant reduction in the effect of the higher two conditions on the academic outcome measures. Dickenson and Truax (1966) explored the impact of time limited group counseling on the achievement levels of underachieving emotionally disturbed college students. Students who received only moderate levels of Accurate Empathy, Warmth, and Genuineness did not differ significantly in their achievement from their controls. However, the children who saw highly facilitative counselors did show significantly greater improvement in their achievement than either of the above two groups. Perhaps relatively higher mean levels of conditions are necessary to effect changes in achievement than are required to alter behavioral or personality dimensions. Further research will hopefully clarify this possibility and the overall impact of the therapeutic conditions on cognitive functions in children.

Rogers, et al. (1967) stressed the importance of the patient's perception of the levels of the conditions present as crucial to the degree of change observed. In essence, Rogers was suggesting that absolutely high levels of conditions can be ineffective if they are not recognized by the client. This could have played an important role in this study, in that no relationship inventory was employed (a measure of the client's perception of the levels of therapist warmth, empathy, and genuineness). Other studies have demonstrated that this measure is often more predictive of the outcome than the process ratings of the conditions (Kurtz and Grummon, 1972).

No assessment of the severity or the etiology of the children's
learning and behavior problems was attempted in this study. Rogers, et al. (1967) found that the severity of the disturbance was related to the level of therapeutic conditions offered. It seems possible that for many of the children tutored, their poor achievement may not have been caused by any emotional (neurotic) learning inhibitions which theoretically would be ameliorable through a therapeutic relationship, but rather could have resulted simply from their poor preparation for academic tasks and their delayed levels of cognitive development. This possibility is further supported by the finding that generally the children with the poorest pre-tutoring scores tended to make the greatest improvement, regardless of tutor and/or the presence of therapeutic conditions. Thus, those who improved the most may have simply been the most deprived, who merely needed a stimulating environment and the appropriate feedback that would enable them to begin to learn previously unfamiliar academic tasks. Baer (1958), Carter (1956, 1957), and Norman, Clarke, and Bessemer (1962) found overplacement to be a major factor in poor and underachievement. Perhaps the youngest children's poor academic performance is merely due to developmental delay, which initially may not require therapeutic interventions to overcome, but rather simply supportive rote practice on academic tasks. Younger children tended to make the greatest improvement on the Iowa Word Analysis and Reading Comprehension GE change scores and on the Bender-Gestalt change score.

Perhaps a more effective experimental design would have included a differential diagnosis of the children's academic problems, one that considers the specific causes of the delayed academic performance and
the degree of associated neurotic learning inhibition. Those children where there is no clear emotional component to their learning disability might then have been referred to another program rather than to therapeutic tutoring, since therapeutic tutoring should be maximally effective with emotionally related underachievement.

This study did not provide a clear indication of the effectiveness of the total training procedures, in that not all the therapeutic principles taught to the tutors were assessed and examined (See Rie, 1974). It is clear however that the content of the sessions was broad, not a delimited school focus, and that the overall level of conditions, especially Accurate Empathy, provided the children were rather low. Further research will possibly clarify the impact of the non-relationship principles to the overall outcome. In addition, every attempt should be made to increase the levels of conditions provided by the tutors in subsequent years.

The nature of the relationship between cognitive and emotional development in children is not yet fully understood nor is the boundary between one's feelings and thoughts. Further, the importance of the therapeutic conditions in producing cognitive changes in disadvantaged children is not yet known. In fact, the remediability of the disadvantaged child's cognitive deficits is currently being debated by therapists. Perhaps, if more traditional therapeutic outcome measures had been employed (personality and intrapsychic measures have been used in many studies and are often considered to be more closely related to the emotional components of the personality that are alterable through a therapeutic experience) a different relationship between the therapeutic
conditions and the outcome measures would have been observed.

The significant correlations between the condition scores and the Classroom Sociogram change scores suggest that facilitative therapists produced some basic personality and behavioral changes in the children. The negative relationship between the conditions and the PTQ-P and PTQ-T (though correlations were small and only significant for the PTQ-P Total score correlated with the combined ratings of Nonpossessive Warmth and Genuineness) suggests the possibility that the children tutored by facilitative therapists changed in ways that were disapproved of by their parents and teachers (i.e., greater spontaneity and expression of feelings). The parents and teachers possibly expected only academic changes in their children and perhaps were not prepared for behavioral or personality changes and consequently responded unfavorably on the post-tutoring PTQ. This possibility is also supported somewhat by the finding that the children tutored by the most highly educated tutors tended to make the greatest improvement on the PTQ-T and also on the PTQ-P Total change scores ($r = .626, p < .01$ for the PTQ-T; $r = .485, p < .05$ for the PTQ-P). Thus, the educational level of the tutors (despite little variation among the group, $S.D. = 1.34$) accounts for 36% of the variance in their ratings on the PTQ-T. This suggests a possible teacher bias that educational level exerts a considerable influence on therapeutic tutoring effectiveness. In addition, most teachers and parents conception and understanding of the therapeutic tutoring program disregarded the program's interest in the child's affective development and conceived of the program only as a traditional remedial tutoring program.
The changes in the children's academic skills that were observed in this study were unrelated to the level of therapeutic conditions offered, and apparently resulted from factors other than tutor empathy, warmth, and genuineness. Basic personality and behavioral changes in the children may have resulted from their involvement with facilitative tutors, but these changes might have gone undetected by the academic measures and potentially biased parent and teacher behavior ratings employed in this study.

Perhaps, the time spent between the pre and post tutoring evaluations of the children was not of sufficient duration to allow the personality or emotional changes in the children to have an impact on their learning potential and their achievement levels. If the children tutored by facilitative tutors did feel better about themselves, their intelligence, and their school as a consequence of the tutoring, the impact of these changes on their academic functioning may require more than five months to be detectable on the traditional achievement measures employed in this study. Carter (1967) found the reading scores of his disadvantaged first graders were not improved in the short run (pre versus post), but were facilitated when assessed for long-term impact twenty months after the cessation of the program. Therefore, perhaps achievement measures are affected by a facilitative relationship, but only over a longer period of time than allowed for in this study.

Many of the psychometric measures employed in this study were selected because of their apparent or purported relationship with and/or their similarity to the therapeutic conditions explored in this study. However, as with the condition data, very few of the psychometric meas-
eras were predictive of the outcome variables. Many studies have been unable to use psychometrically derived data to predict outcomes (effectiveness) in teaching, counseling, or therapy (Williams, et al., 1966; Schluck, 1970; Openshaw, 1967; Watton, 1970; and Christensen, 1960). Allen (1967) concluded that no therapist personality characteristics were predictive of therapeutic effectiveness. Fulkerson and Barry (1961) and Luborsky, et al. (1971) concluded after a thorough review of the literature that therapist and client variables having the strongest relationship to the outcome have been non-test variables. Clearly, relatively few of our traditional psychometric measures achieve a persistently high predictive validity and the field awaits the refinements that will enable us to make more reliable long term predictions.

Considering the findings cited above, it is not surprising that very few of the psychometric measures employed in this study were ineffective predictors of tutoring outcome. The tutor dogmatism scores were unrelated to the outcome change scores. Petty (1970) found that dogmatism was not predictive of counselor trainee success. Parsons and Olsen (1969), Stoffer (1968), and Foulds (1971) found that dogmatism scores did not discriminate facilitative from non-facilitative beginning counselors and helpers.

Similarly, internal tutors were not found to be more effective. Rotter (1966) noting the failure of the I-E scale to obtain strong relationships with a multitude of variables suggested that the presence of situationally specific attitudes in the scale and in the behaviors being investigated may be at fault (p.21). Regardless, locus of control was found to be an ineffective predictor of therapeutic
Generally, the TSCS was also found to be an ineffective predictor of the outcome variables (except for P Score Identity with PTQ-T Total change score and the Total V Score with the Iowa Reading Comprehension GE change score). The overall low correlations of the TSCS subtests with the outcome scores is not surprising since the nature of this scale would contraindicate very many high linear correlations. Rather, only the high and low scores for the variables might be expected to be correlated. Thus, a generalized measure of correlation (i.e., a correlation ratio) might have yielded different results if it had been employed in this study.

The tutor's intelligence scores were unrelated to their effectiveness and in addition their levels of test derived empathy were found to be unrelated to their children's improvement rates on each of the outcome measures.

Very few of the tutor psychometric measures were predictive of the level of rated therapeutic conditions. This was somewhat surprising because several of the measures were selected because of their apparent relationship to the condition ratings. For example, the Self-Criticism subscale of the TSCS is considered to measure ones openness, defensiveness, and capacity for criticism. Characteristics that would seemingly be related to the levels of rated genuineness. However, although in the predicted direction, this correlation was insignificant (r = .110, p > .05).

In like manner, a paper and pencil measure of empathy (the Hogan Empathy Scale) was insignificantly related to process ratings of
Accurate Empathy ($r = .229, p > .05$). This suggests that either the rating procedure was invalid or that these two measures of empathy actually assess different characteristics of the tutors, both of which are called empathy. Kurtz and Grummon (1972) found that six different measures of empathy were unrelated to each other. Further, they found that only process ratings of empathy were related to the depth of the client's self-exploration. In addition, only the process rated empathy and the client's perceptions of therapist empathy were related to the study's outcome measures.

The findings that the Personality Integration and Psychoses subscales of the TSCE are predictive of the rated condition levels (see Table 20) indicates simply that the better adjusted tutors were rated as warmer, more genuine, and higher in combined condition levels, than were the less well adjusted tutors. This could reflect the biases of the raters or it could reflect a valid and reliable characteristic of the tutors. However, since many studies have found that the degree of personality integration of the therapist is highly related to his therapeutic effectiveness the latter conclusion seems more justified.

CRITICISMS OF THIS STUDY

A number of the weaknesses of this study were already outlined or alluded to in previous sections and consequently will not be reconsidered here. The following are some other outstanding weaknesses in this field study.

1. This study employed change scores (post-pre tutoring) as its only measure of therapeutic outcome. Perhaps, these change scores
should have been converted to standard scores and then to t scores, which would have allowed for the calculation of a composite outcome variable(s). In addition, the assumption that all changes of equal magnitude on the outcome measures reflects equivalent improvement remains to be documented.

2. Only two raters were employed in this study, the author and his wife. Three raters would have been more desirable, with nine being ideal. In that case, three different people could have rated each condition. These experimental procedures simply were not feasible for this study.

3. Only one session was recorded by the tutors. Consequently, novelty effects, etc., were not controlled for. A dummy or practice taping would have been desirable and probably would have precluded the loss of subjects resulting from inaudible recordings.

4. Only late therapeutic tutoring sessions were rated and although the literature suggests that this is the optimal procedure, there are no data from which to determine the inter-session condition variability for minimally trained lay therapeutic tutors. Thus, the basic question of when during the tutoring process did the process movement occur, remains unanswered. Perhaps early and middle tutoring sessions should have been recorded to help clarify this important issue.

5. This study did not control for initial differences in the children's level of achievement motivation. This variable could be assumed to be random, but since the tutors tutored different numbers of children from different schools (with some variation among the schools
in severity of student academic disability) perhaps this potential source of variance should have been considered and controlled for.

**IMPLICATIONS FOR FURTHER RESEARCH**

Further research is needed which will clarify, support, or disprove the tentative findings of this study. The following additional suggestions are offered.

1. Further research should include more specifically diagnosed groups, and relate the process ratings of the therapeutic conditions offered to a wider variety of outcome measures.

2. A large sampling of audio tapes from the entire tutoring period should be analyzed to determine the inter-session stability of therapeutic conditions for the therapeutic tutors.

3. The importance of a therapeutic relationship in the amelioration of culturally related poor and underachievement is yet to be demonstrated. The variables observed in this study account for very little of the total variance. Perhaps further research will clarify these issues.

4. Future studies may systematically manipulate the level of therapeutic conditions offered the child, which would allow for causal interpretations that cannot be made with correlational studies.

5. The findings of this study suggest the possibility that non-therapist factors may play a significant role in effecting changes during therapeutic tutoring. Future studies might investigate the contribution to the tutoring outcome of specific relationship characteristics and/or child variables that might be related to the level of
therapeutic conditions offered and to the overall tutoring outcome.

6. The PTQ-T data could be factor analyzed with the resulting factors then being related to the sociogram ratings in an attempt to determine more specifically the meaning of the Classroom Sociogram total score and exactly what it is that the children are responding to in their selections on this scale.
APPENDIX A

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 with slight accuracy towards the client's obvious feelings. He shows concern with the deeper feelings but is inaccurate with regard to them.

*Bergin and Solomon's 1970 revision (Stage 3 is added) of Truax's original scale.
STAGE 4

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 client.

STAGE 5

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 6

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 6, the therapist simply communicates his awareness of the problem of understanding another person's inner world. Stage 6 is the midpoint of the continuum of accurate empathy.

STAGE 7

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 8 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 8

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 8 is distinguishable from Stage 7 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 9

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 re­
sulting minor flaws in the accuracy of his understanding, but inaccura­
cies 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
9 makes mistakes, mistakes do not have a jarring note, but are covered
by the tentative character of the response. Also the therapist is sensi­
tive to his mistakes and quickly alters or changes his response in mid­
stream, 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.
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 (Truax and Carkhuff, 1967, pp. 47-57).
APPENDIX B

A SCALE FOR THE MEASUREMENT OF NON-POSSESSIVE WARMTH

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 personal 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 to 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 (Truax and Carkhuff, 1967, pp. 60–67).
APPENDIX C
A SCALE FOR THE MEASUREMENT OF GENUINENESS

Charles B. Truax

STAGE 1

The therapist is clearly defensive in the interaction and there is explicit evidence of a very considerable discrepancy between what he says and what he experiences. There may be striking contradictions in the therapist's statements. The content of his verbalizations may contradict the voice qualities or non-verbal cues.

STAGE 2

The therapist responds appropriately but in a professional, rather than a personal manner, giving the impression that his responses are said, because they sound good from a distance, but do not express what he really feels or means. There is a somewhat contrived or rehearsed quality or air of professionalism present.

STAGE 3

The therapist is implicitly either defensive or professional although there is no explicit evidence.

STAGE 4

There is neither implicit nor explicit evidence of defensiveness or the presence of a facade. The therapist shows no self-incongruence.

STAGE 5

The therapist is freely and deeply himself in the relationship. He is open to experiences and feelings of all types, both pleasant and hurtful, without traces of defensiveness or retreat into profession-
Although there may be contradictory feelings, these are accepted or recognized. The therapist is clearly being himself in all of his responses, whether they are personally meaningful or trite. At Stage 5 the therapist need not express personal feelings, for whether he is giving advice, reflecting, interpreting, or sharing experiences, it is clear that he is being very much himself so that his verbalizations match his inner experiences (Truax and Carkhuff, 1967, pp. 68-72).
APPENDIX D

PARENT-TEACHER QUESTIONNAIRE

Please rate items based on the child's present behavior. Rate every item independently. Your first impressions are desirable.

USE THE RATING SCALE BELOW. WRITE YOUR RATING (NUMBER) FOR EACH ITEM ON THE BLANK TO THE LEFT OF THE ITEM NUMBER. PLEASE ANSWER EVERY ITEM EVEN IF YOU NEED TO GUESS.

<table>
<thead>
<tr>
<th>Very frequently</th>
<th>Often</th>
<th>Occasionally</th>
<th>Rarely</th>
<th>Never</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>

COMpared to the average child, how often does the child... . . . . . .

1. sit still when he needs to wait his turn?
2. answer quickly and correctly when called on or questioned?
3. come to you for help in doing things when necessary?
4. stick to an activity he selected for himself for fifteen minutes or more?
5. stick to an activity without jumping from one thing to another?
6. refrain from pulling, pushing, or grabbing other children or adults?
7. behave well when he is with others in a group?
8. complete his work assignments?
9. respond in a friendly fashion to other children and adults?
10. maintain attention when the teacher and/or parent is explaining something to him?
11. stick to one subject while talking?
12. refrain from bothering other people who are obviously busy?
13. stay in his seat during lunch or dinner?
14. show continuous improvement in his reading, writing, and arithmetic skills?
15. seek out other children to play with?
16. stick to the point when he is in conversation?
17. tolerate noises (classroom or playground noises) or movement?
18. refrain from hitting or hurting other children?
19. refrain from fidgeting or moving around unnecessarily?
20. stick to something when it is somewhat difficult or frustrating?
21. play with or relate to adults?
22. look alert when you are having a conversation with him?
23. continue his work or task when being annoyed by others?
24. refrain from disrupting others by making noises, and faces?
25. exhibit good control over his body movements?
26. get the point of what he hears or reads?
27. describe interesting things that have happened to him?
28. give a relevant answer to a question?
29. refrain from whining, crying, or complaining?
30. stick to his goal or intention to do something?
31. read or look at picture books by himself?
32. watch an interesting movie or television show with good attention?
33. learn quickly when interested in the activity?
34. want to learn how to make things or do things?
35. make things, color, or write letters on his own without being asked to do so?
APPENDIX E

CLASSROOM SOCIОGRAM

Child's Name:__________________________

Teacher:

Please ask the children in the classroom to indicate on paper (not publicly) the three children they would most like to sit next to. If the children have difficulty writing names, you may interview them individually.

When they have listed the names, and they are in your possession, count the number of times the study child has been chosen, first, second and third. Individual preferences must be made. Then fill out the blanks below.

1) Number of times child was chosen first___________
2) Number of times child was chosen second___________
3) Number of times child was chosen third___________
4) Number of boys in your classroom________________
5) Number of girls in your classroom________________
6) Number of children in your classroom______________

Thank you for your help!
APPENDIX F

ROCKEFELLER DOGMATISM SCALE

NAME: ___________________________ DATE: ___________________________

Opinions and Reactions Inventory 1.

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.

Please mark each statement at the right, according to how much you agree or disagree with it. Please mark every item. Write 1, 2, 3, or -1, -2, -3, depending on how you feel in each case. Plus means you agree with the statement, minus means you disagree with the statement as follows:

+1 I agree slightly       -1 I disagree slightly
+2 I agree moderately    -2 I disagree moderately
+3 I agree strongly       -3 I disagree strongly

1. A person who thinks primarily of his own happiness is beneath contempt. ____________________________________ 1.

2. The main thing in life is for a person to want to do something important. ____________________________________ 2.

3. In a discussion I often find it necessary to repeat myself several times to make sure I am being understood. ____________________________________ 3.

4. Most people just don't know what's good for them. ____________________________________ 4.

5. In times like these, a person must be pretty selfish if he considers his own happiness primarily. ____________________________________ 5.
6. A man who does not believe in some great cause has not really lived.

7. I'd like it if I should find someone who would tell me how to solve my personal problems.

8. Of all the different philosophies which have existed in this world there is probably only one which is correct.

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

10. 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.

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

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

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

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

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

16. 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.

17. Even though freedom of speech for all groups is a worthwhile goal, it is unfortunately necessary at times to restrict the freedom of certain political groups.

18. If a man is to accomplish his mission in life it is sometimes necessary to gamble "all or nothing at all."

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

20. A person who gets enthusiastic about a number of causes is likely to be a pretty wishy-washy sort of person.

21. To compromise with our political opponents is dangerous because it usually leads to the betrayal of our own side.
22. If given the chance I would do something that would be of great benefit to the world.

23. 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.

24. 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.

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

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

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

28. The United States and Russia have just about nothing in common.

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

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

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

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

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

34. 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.

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

36. 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.
37. Most of the ideas which get published nowadays aren't worth the paper they are printed on.

38. It is only natural for a person to be rather fearful of the future.

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

40. 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.
APPENDIX C

ROTTER'S I-E

Opinions and Reactions Inventory 2.

This survey is designed to measure people's opinions. For each question, circle the alternative (either "a" or "b") which most nearly represents your own views. For some questions you may agree or disagree with both alternatives but in either case, circle the one with which you are in closest agreement. Be sure and answer every item. There are no right or wrong answers and you can be sure that many others feel as you do!

1. a. Children get into trouble because their parents punish them too much.
   b. The trouble with most children nowadays is that their parents are too easy with them.

2. a. Many of the unhappy things in people's lives are partly due to bad luck.
   b. People's misfortunes result from the mistakes they make.

3. a. One of the major reasons why we have wars is because people don't take enough interest in politics.
   b. There will always be wars, no matter how hard people try to prevent them.

4. a. In the long run people get the respect they deserve in this world.
   b. Unfortunately, an individual's worth often passes unrecognized no matter how hard he tries.

5. a. The idea that teachers are unfair to students is nonsense.
   b. Most students don't realize the extent to which their grades are influenced by accidental happenings.

6. a. Without the right breaks one cannot be an effective leader.
   b. Capable people who fail to become leaders have not taken advantage of their opportunities.

7. a. No matter how hard you try some people just don't like you.
   b. People who can't get others to like them don't understand how to get along with others.

8. a. Heredity plays the major role in determining one's personality.
   b. It is one's experiences in life which determine what they're like.
9. a. I have often found that what is going to happen will happen.
b. Trusting to fate has never turned out as well for me as making a decision to take a definite course of action.

10. a. In the case of the well prepared student there is rarely if ever such a thing as an unfair test.
b. Many times exam questions tend to be so unrelated to course work that studying is really useless.

11. a. Becoming a success is a matter of hard work, luck has little or nothing to do with it.
b. Getting a good job depends mainly on being in the right place at the right time.

12. a. The average citizen can have an influence in government decisions.
b. This world is run by the few people in power, and there is not much the little guy can do about it.

13. a. When I make plans, I am almost certain that I can make them work.
b. It is not always wise to plan too far ahead because many things turn out to be a matter of good or bad fortune anyhow.

14. a. There are certain people who are just no good.
b. There is some good in everybody.

15. a. In my case getting what I want has little or nothing to do with luck.
b. Many times we might just as well decide what to do by flipping a coin.

16. a. Who gets to be the boss often depends on who was lucky enough to be in the right place first.
b. Getting people to do the right thing depends upon ability, luck has little or nothing to do with it.

17. a. As far as world affairs are concerned, most of us are the victims of forces we can neither understand, nor control.
b. By taking an active part in political and social affairs the people can control world events.

18. a. Most people don’t realize the extent to which their lives are controlled by accidental happenings.
b. There really is no such thing as "luck".

19. a. One should always be willing to admit mistakes.
b. It is usually best to cover up one’s mistakes.

20. a. It is hard to know whether or not a person really likes you.
b. How many friends you have depends upon how nice a person you are.
21. a. In the long run the bad things that happen to us are balanced by the good ones.
b. Most misfortunes are the result of lack of ability, ignorance, laziness, or all three.

22. a. With enough effort we can wipe out political corruption.
b. It is difficult for people to have much control over the things politicians do in office.

23. a. Sometimes I can't understand how teachers arrive at the grades they give.
b. There is a direct connection between how hard I study and the grades I get.

24. a. A good leader expects people to decide for themselves what they should do.
b. A good leader makes it clear to everybody what their jobs are.

25. a. Many times I feel that I have little influence over the things that happen to me.
b. It is impossible for me to believe that chance or luck plays an important role in my life.

26. a. People are lonely because they don't try to be friendly.
b. There's not much use in trying too hard to please people, if they like you, they like you.

27. a. There is too much emphasis on athletics in high school.
b. Team sports are an excellent way to build character.

28. a. What happens to me is my own doing.
b. Sometimes I feel that I don't have enough control over the direction my life is taking.

29. a. Most of the time I can't understand why politicians behave the way they do.
b. In the long run the people are responsible for bad government on a national as well as on a local level.
APPENDIX H

HOGAN EMPATHY SCALE

Personal Reaction Inventory

This inventory contains a series of statements. Read each one carefully. If you agree with a statement and decide it is true about you, answer TRUE (T). If you disagree with a statement, and feel that it is not true about you, answer FALSE (F). Please mark your answer on the booklet by circling your preferred response.

If you find a few questions which you cannot or prefer not to answer, they may be omitted. However, it is important that you give your own honest opinion about yourself and that you try to answer every item on the inventory. Please erase completely any answer you wish to change and be sure you have written your name in the space provided.

NAME: ___________________   DATE: ______________

T F 1. A person needs to "show off" a little now and then.

T F 2. I liked "Alice in Wonderland" by Lewis Carroll.

T F 3. Clever, sarcastic people make me feel very uncomfortable.

T F 4. I usually take an active part in the entertainment at parties.

T F 5. I feel sure that there is only one true religion.

T F 6. I am afraid of deep water.

T F 7. I must admit I often try to get my own way regardless of what others may want.

T F 8. I have at one time or other in my life tried my hand at writing poetry.

T F 9. Most of the arguments or quarrels I get into are matters of principle.

T F 10. I would like the job of a foreign correspondent for a newspaper.
T F 11. People today have forgotten how to feel properly ashamed of themselves.

T F 12. I prefer a shower to a bathtub.

T F 13. I always try to consider the other fellow's feelings before I do something.

T F 14. I usually don't like to talk much unless I am with people I know very well.

T F 15. I can remember "playing sick" to get out of something.

T F 16. I like to keep people guessing what I am going to do next.

T F 17. Before I do something I try to consider how my friends will react to it.

T F 18. I like to talk before groups of people.

T F 19. When a man is with a woman he is usually thinking about things related to her sex.

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

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

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

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

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

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

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

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

T F 28. I have a natural talent for influencing people.

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

T F 30. I feel like giving up quickly when things go wrong.

T F 31. It is hard for me just to sit still and relax.
F 32. Once in a while I think of things too bad to talk about.
F 33. I feel that it is certainly best to keep my mouth shut when I'm in trouble.
T F 34. I am a good mixer.
T F 35. I am an important person.
T F 36. I like poetry.
T F 37. My feelings are not easily hurt.
T F 38. I have met problems so full of possibilities that I have been unable to make up my mind about them.
T F 39. Often I can't understand why I have been so cross and grouchy.
T F 40. What others think of me does not bother me.
T F 41. I would like to be a journalist.
T F 42. I like to talk about sex.
T F 43. My way of doing things is apt to be misunderstood by others.
T F 44. Sometimes without any reason or even when things are going wrong I feel excitedly happy, "on top of the world."
T F 45. I like to be with a crowd who play jokes on one another.
T F 46. My mother or father often made me obey even when I thought that it was unreasonable.
T F 47. I easily become impatient with people.
T F 48. Sometimes I enjoy hurting persons I love.
T F 49. I tend to be interested in several different hobbies rather than to stick to one of them for a long time.
T F 50. I am not easily angered.
T F 51. People have often misunderstood my intentions when I was trying to put them right and be helpful.
T F 52. I am usually calm and not easily upset.
T F 53. I would certainly enjoy beating a crook at his own game.
T F 54. I am often so annoyed when someone tries to get ahead of me in a line of people that I speak to him about it.
55. I used to like hopscotch.
56. I have never been made especially nervous over trouble that any members of my family have gotten into.
57. As a rule I have little difficulty in "putting myself into other people's shoes.
58. I have seen some things so sad that I almost felt like crying.
59. Disobedience to the government is never justified.
60. It is the duty of a citizen to support his country, right or wrong.
61. I am usually rather short-tempered with people who come around and bother me with foolish questions.
62. I have a pretty clear idea of what I would try to impart to my students if I were a teacher.
63. I enjoy the company of strong-willed people.
64. I frequently undertake more than I can accomplish.
Taping of Therapeutic Tutoring Sessions

Name of Tutor:_____________________________ Date:____________________

Name of Child:____________________________________________________________________

After completing the above information please rate on the scale below the degree to which this session, which you taped, was like other sessions you have previously experienced with this particular child. (Circle only one). Please complete one form for each child you tutor.

1 2 3 4

Very Similar  Similar  Different  Very Different

If you circled either three or four, please describe below, as specifically as you can, how this session was different from your normal session with this particular child. Thank you very much!
APPENDIX J

THERAPEUTIC TUTORING PROGRAM 1972-1973

Section on Child Development and Psychology

Children’s Hospital

AUDIO TAPING OF THERAPEUTIC TUTORING SESSIONS

As we approach the conclusion of the first year of the therapeutic tutoring project we are faced with the important task of evaluating the effectiveness of the program. Often, at an early point in the development of an educational program many approaches are undertaken to evaluate the program's merit and to discover ways in which the program can be improved. We are very pleased with the progress of the program this year and we realize that each of you has acquired valuable experience during the course of the year. Therefore we are asking each tutor to record one audio tape of a session with each child she is currently tutoring. Through the audio taping of your sessions we hope to learn what the children are doing, how they respond during the sessions, and what elements or characteristics seem to foster the greatest changes in their academic performance and behavior. We also hope to gain some insights into the difficult and perplexing question of why some children make greater gains than others. Analyses of these tapes will also enable us to improve the effectiveness of the training program by making us more aware of the behavior styles and response patterns of the children being referred for tutoring and also allow us to use actual tapes of
We realize that each of you has a busy schedule, but we feel that the taping of individual sessions is an important part of this program and hope it is not a major imposition on your time. We hope that you will do exactly what you typically do with the children. Please **DO NOT** change your procedure because you are being taped. We are interested in observing your **routine** session. Since we are interested in the range of **variability** of procedures being employed by the various tutors with the children, creative or atypical events would also be of interest and could be taped. The important point is that over all, the tapes be representative of what our therapeutic tutors routinely do with their children.

The tutor should explain the purpose of the taping honestly and briefly to the child. She should explain that the tapes will allow for the directors of the project at the Children's Hospital, to determine how the project has worked and how it can be improved in the future. If the child seems interested in the tape, some portion of the tape could be replayed for him at the end of the session.

The tutor should set the recorder in a convenient place and be sure that it is recording properly before starting the session. Since children often mumble, the volume should be turned up and the microphone appropriately directed toward the child. Please be sure you speak loudly and clearly and it is a good idea to check the record level needle at least once. The entire session, for each child, should be taped. After completing each taping please rewind it and check the tape briefly to see that it recorded properly. Before starting your session with the child be sure to state on the tape itself your name, the date, and the
name of the child being tutored. Obviously this is extremely important. In addition, please write the names of the appropriate children on each tape or on the box. After recording a session please complete the attached forms (one should be completed for each child) describing how typical the taped session was comparatively to other sessions with that particular child.

Most of the tape recorders you will be using are borrowed and consequently we urge you to please handle them with care and to be extremely careful that they are not stolen. We place them confidently in your responsibility.

As soon as you have completed the taping of one session with each child you are currently tutoring, please call Mrs. Deborah Henry and inform her you are finished. She can be reached at Children's Hospital at 253-8841, ext. 466. Since we have to share to recorders between several tutors we encourage you to begin immediately and consequently no one should require more than a week to complete the tapings.

If there are any problems, questions, etc., please contact me as soon as possible at 253-8841, ext. 466, or at 866-0671.

Thank you for your cooperation in this mutual endeavor.

Sincerely,

Robert T. Young,
Project Assistant
### APPENDIX K

MEANS AND STANDARD DEVIATIONS FOR THE NORMATIVE POPULATIONS OF THE TENNESSEE SELF-CONCEPT SCALE, ROKEACH'S DOGMATISM SCALE, ROTTER'S I-E, AND HOGAN'S EMPATHY SCALE

<table>
<thead>
<tr>
<th>Scale</th>
<th>Mean</th>
<th>Standard Deviation</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>TENNESSEE SELF-CONCEPT SCALE</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Self Criticism</td>
<td>35.54</td>
<td>6.70</td>
<td>626</td>
</tr>
<tr>
<td>Total P Score</td>
<td>345.57</td>
<td>30.70</td>
<td></td>
</tr>
<tr>
<td>P Score Identity</td>
<td>127.10</td>
<td>9.96</td>
<td></td>
</tr>
<tr>
<td>Total V Score</td>
<td>48.53</td>
<td>12.42</td>
<td></td>
</tr>
<tr>
<td>Personality</td>
<td>10.42</td>
<td>3.88</td>
<td></td>
</tr>
<tr>
<td>Integration</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rokeach Dogmatism</td>
<td>132.2*</td>
<td>22.5</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>130.6**</td>
<td>-----</td>
<td>30</td>
</tr>
<tr>
<td>Rotter's I-E</td>
<td>9.05***</td>
<td>3.66</td>
<td>116</td>
</tr>
<tr>
<td>Hogan Empathy</td>
<td>40.7X</td>
<td>5.10</td>
<td>143</td>
</tr>
</tbody>
</table>

* Female College Students (Hogan, 1969)
* American teachers in summer school (Rabkin, 1966)
** NDEA Guidance and Counseling Institute (Saltzman, 1966)
*** Negro Psychology Students, Florida State University (Gore and Rotter, 1963)
BIBLIOGRAPHY


Betz, B. and Whitehorn, J.C. The Relationship of the therapist to the outcome of the therapy in schizophrenia. Psychiatric Research Reports, 1956, 5, 89-105.


Campbell, D.P. Achievements of counseled and non-counseled students twenty-five years after counseling. *Journal of Counseling Psychology*, 1965, 12, 287-293.


Denker, P.G. Results of treatment of psychoneuroses by the general practitioner. New York State Journal of Medicine, 1946, 46, 2164-66.


Ellis, A. The Effectiveness of psychotherapy with individuals who have severe, homosexual problems. Journal of Consulting Psychology, 1956, 20, 191-195.


Fiedler, F.E. The Psychological distance dimension in interpersonal relations. Journal of Personality, 1953, 22, 142-150.


Fulkerson, S.C., and Barry, J.R. Methodology and research on the prognostic use of psychological tests. Psychological Bulletin, 1961, 58, 177-204.


Harvey, L. The Use of nonprofessional auxiliary counselors in staffing a counseling service. Journal of Counseling Psychology, 1964, 11, 348-351.


Harvey, L. The Use of nonprofessional auxilliary counselors in staffing a counseling service. Journal of Counseling Psychology, 1964, 11, 348-351.


Morgan, R.F., and Toy, T.B. Learning by teaching: A Student to student compensatory tutoring program in a rural school system and its relevance to the educational cooperative. Psychological Record, 1970, 20, 159-169.


Norman, R.D., Clark, B.P., and Bessemer, D.W. Age, sex, IQ, and achievement patterns in achieving and non-achieving gifted children. Exceptional Children, 1962, 29, 116-123.


Rogers, C. The Interpersonal relationship in the facilitation of learning. Lecture at Harvard University, April, 12, 1966.


Sayegh, Y., and Dennis, W. The Effect of supplementary experiences upon the behavioral development of infants in institutions. *Child Development*, 1965, 36, 81-90.


Shaw, M.C., and Dutton, B.E. The Use of the Parent Attitude Research Inventory with the parents of bright academic underachievers. *Journal of Educational Psychology*, 1962, 53, 203-308.


Spotts, J. The Perception of positive regard by relatively successful and relatively unsuccessful clients. Brief Research Reports, Wisconsin Psychiatric Institute, University of Wisconsin, 1962, 15.


Truax, C.B. Comparisons between control patients, therapy patients perceiving high conditions and therapy patients perceiving low conditions on measures of constructive personality change. Brief Research Reports, Wisconsin Psychiatric Institute, University of Wisconsin, 1962(a), 31.


Uhlenhuth, E.H., and Duncan, D.B. Subjective change in psychoneurotic outpatients with medical students I, the kind, amount, and course of change. Unpublished Manuscript, John Hopkins University, 1968.


