The Effect of Diagnostic Labeling on the Clinician's Perception of Behavioral Abnormalities

Clifton Joseph Saper
Loyola University Chicago

Follow this and additional works at: https://ecommons.luc.edu/luc_diss

Part of the Psychology Commons

Recommended Citation
https://ecommons.luc.edu/luc_diss/1697

This Dissertation is brought to you for free and open access by the Theses and Dissertations at Loyola eCommons. It has been accepted for inclusion in Dissertations by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.

This work is licensed under a Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License.
Copyright © 1976 Clifton Joseph Saper
THE EFFECT OF DIAGNOSTIC LABELING ON THE CLINICIAN'S
PERCEPTION OF BEHAVIORAL ABNORMALITIES

by

Clifton J. Saper

A Dissertation Submitted to the Faculty of the
Graduate School of Loyola University of
Chicago in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

September
1976
ACKNOWLEDGMENTS

The author wishes to thank James Johnson, Ph.D., Thomas Petzel, Ph.D., and Patricia Barger, Ph.D. who encouraged this study and offered guidance and support whenever it was needed. Appreciation is also extended to John Shack, Ph.D., John Carney, Sylvana Yovanof, and Christine Bowman who assisted at various stages during the project.

To a beloved little friend, Kevin, who planted the seed for this research, gave it life, and hopefully will benefit from the fruit of its application, a great debt of gratitude is owed.
VITA

The author, Clifton Joseph Saper, is the son of Dr. Bernard Saper and Mrs. Johanna (Franklin) Saper. He was born June 21, 1950, in Los Angeles, California.

His elementary education was obtained in the public schools of Chicago, Illinois, and Albany, New York. He was graduated from Bethlehem Central Senior High School located in Delmar, New York, in 1968.

In September, 1968, he entered Brown University in Providence, Rhode Island. In June, 1972, he received the degree of Bachelor of Arts with a major in psychology. While attending Brown University, he was teaching assistant to Dr. Francis Clayton of the Psychology Department.

In September, 1972, he entered the Clinical Psychology doctoral program at Loyola University of Chicago. He was awarded a National Institute of Mental Health training stipend. From June, 1973 to September, 1973, he completed a 500 hour clerkship in clinical psychology at the Bangor Mental Health Institute in Bangor, Maine. From September, 1973 to January, 1975 he completed a 2,000 hour clinical child psychology internship at the Loyola University Guidance Center and Day School.

He recently completed a two-year internship in mental health administration with the Loyola University Guidance
Center. For the year 1975-1976, he was awarded a Loyola University Research Fellowship to undertake the present investigation. Currently, he is a clinical psychologist on an adolescent inpatient unit in Chicago and is a lecturer in psychology at Loyola University.

He is a member of the American Psychological Association and the Association for the Advancement of Behavior Therapy, and has been involved in the Illinois Association of Community Mental Health Agencies and the Illinois Group Psychotherapy Society.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>ii</td>
</tr>
<tr>
<td>VITA</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>viii</td>
</tr>
<tr>
<td>Chapter</td>
<td></td>
</tr>
<tr>
<td>I. INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II. REVIEW OF RELATED LITERATURE</td>
<td>6</td>
</tr>
<tr>
<td>Nature and Function of Diagnoses</td>
<td>6</td>
</tr>
<tr>
<td>Social Perspective</td>
<td>11</td>
</tr>
<tr>
<td>Person Perception and Expectancies</td>
<td>22</td>
</tr>
<tr>
<td>The Nature of Behavioral Observation and Rating</td>
<td>30</td>
</tr>
<tr>
<td>Clinician Variables Which Affect Behavioral Observation, Rating, and the Therapeutic Relationship</td>
<td>34</td>
</tr>
<tr>
<td>Methodological Issues</td>
<td>45</td>
</tr>
<tr>
<td>Pilot Study</td>
<td>51</td>
</tr>
<tr>
<td>Hypotheses</td>
<td>53</td>
</tr>
<tr>
<td>III. METHOD</td>
<td>55</td>
</tr>
<tr>
<td>Subjects</td>
<td>55</td>
</tr>
<tr>
<td>Sorting subjects into categories</td>
<td>56</td>
</tr>
<tr>
<td>Undergraduate sample</td>
<td>59</td>
</tr>
<tr>
<td>Materials</td>
<td>59</td>
</tr>
<tr>
<td>The films</td>
<td>59</td>
</tr>
<tr>
<td>Clinician Questionnaire</td>
<td>61</td>
</tr>
<tr>
<td>Behavior Description Test</td>
<td>63</td>
</tr>
<tr>
<td>Semantic differential</td>
<td>65</td>
</tr>
<tr>
<td>Peterson Problem Checklist</td>
<td>67</td>
</tr>
<tr>
<td>Procedure</td>
<td>68</td>
</tr>
<tr>
<td>IV. RESULTS</td>
<td>75</td>
</tr>
<tr>
<td>The Effect of Diagnostic Labels on Observers' Interpretation of Behavior</td>
<td>76</td>
</tr>
<tr>
<td>Secondary Concerns</td>
<td>83</td>
</tr>
</tbody>
</table>
### TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>V. DISCUSSION</strong></td>
<td>89</td>
</tr>
<tr>
<td>Labeling Bias: General</td>
<td>89</td>
</tr>
<tr>
<td>Perceived Pathology: A Function of Training and Actual Pathology</td>
<td>91</td>
</tr>
<tr>
<td>&quot;Traditional&quot; vs. &quot;Behavioral&quot; vs. Undergraduate Observers</td>
<td>95</td>
</tr>
<tr>
<td>Secondary Concerns</td>
<td>99</td>
</tr>
<tr>
<td>Methodological Issues</td>
<td>105</td>
</tr>
<tr>
<td>Future Research</td>
<td>116</td>
</tr>
<tr>
<td>Implications</td>
<td>118</td>
</tr>
<tr>
<td><strong>REFERENCES</strong></td>
<td>124</td>
</tr>
<tr>
<td><strong>REFERENCE NOTES</strong></td>
<td>132</td>
</tr>
<tr>
<td><strong>Appendices</strong></td>
<td>133</td>
</tr>
<tr>
<td><strong>A. LETTER REQUESTING THERAPISTS' PARTICIPATION</strong></td>
<td>133</td>
</tr>
<tr>
<td><strong>B. CLINICIAN QUESTIONNAIRE</strong></td>
<td>135</td>
</tr>
<tr>
<td><strong>C. LETTER EXPLAINING RESEARCH TO PARTICIPANTS</strong></td>
<td>138</td>
</tr>
<tr>
<td><strong>D. DESCRIPTION OF CHILDREN IN FILMS</strong></td>
<td>140</td>
</tr>
<tr>
<td>I. Description of Children in Films by a Professional Developmental Psychologist</td>
<td>141</td>
</tr>
<tr>
<td>II. Description of Children in Films by an English Teacher</td>
<td>142</td>
</tr>
<tr>
<td>III. Description of Children in Films by an Untrained Undergraduate Psychology Student</td>
<td>143</td>
</tr>
<tr>
<td><strong>E. BEHAVIOR DESCRIPTION TEST WITH KEY</strong></td>
<td>144</td>
</tr>
<tr>
<td><strong>F. SEMANTIC DIFFERENTIAL WITH KEY</strong></td>
<td>147</td>
</tr>
<tr>
<td><strong>G. PETERSON PROBLEM CHECKLIST</strong></td>
<td>151</td>
</tr>
</tbody>
</table>
### TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Appendices (continued)</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>H. COVER SHEETS FOR TEST BOOKLETS</strong></td>
<td>154</td>
</tr>
<tr>
<td>I. Normal Girl Description Sheet</td>
<td>155</td>
</tr>
<tr>
<td>II. Normal Boy Description Sheet</td>
<td>156</td>
</tr>
<tr>
<td>III. Disturbed Girl Description Sheet</td>
<td>157</td>
</tr>
<tr>
<td>IV. Disturbed Boy Description Sheet</td>
<td>158</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Demographic Data on the Trained Clinician Observer Sample</td>
<td>58</td>
</tr>
<tr>
<td>2. Means and Standard Deviations of the Transformed Ratings on All Four Measures for All Experimental Groups</td>
<td>77</td>
</tr>
<tr>
<td>3. Analysis of Variance Summary Table for Combined Ratings on All Four Measures</td>
<td>80</td>
</tr>
<tr>
<td>4. Analyses of Variance F Values for Each Measure Individually</td>
<td>85</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

Most professionals in the mental health field would agree that conventional psychiatric diagnoses may be useful in clinical practice (e.g., Caveny, Wittson, Hunt & Herrman, 1955; Gough, 1971; Klopfer, 1962; Zigler & Phillips, 1961; Zubin, 1967). However, numerous studies have found diagnostic labels to be statistically unreliable (e.g., Braginsky & Braginsky, 1974; Edelman, 1969; Koestler, 1975; Rosenhan, 1973; Taft, 1955; Temerlin, 1968; Yates, 1970). In some of these studies clinicians were found to diagnose the same patient with very different labels (Temerlin, 1968). In other studies, a patient's characteristics were not always correctly predicted from his/her psychiatric label (Phillips, 1963).

The investigation of the problems surrounding clinical labeling necessarily draws together research from three disparate, but often overlapping, areas in the field of psychology. The first area relates to the issue of the true nature and utility of conventional psychiatric diagnoses and clinical judgment (some salient examples are: Berdie, 1950; Caveny et al., 1955; Chein, 1966; Gough, 1971; Hobbs, 1975; Klopfer, 1962; Meehl, 1956; Winthrop, 1964; Zigler & Phillips, 1961; Zubin, 1967).
1961). A second area of concern deals with the biases inherent in clinical observation and the utilization of various social and behavioral cues (e.g., Barker, 1951; Bersoff, 1971; Goldstein, 1962; Gustin, 1969; Price, 1973; Rabkin, 1972; Rosenhan, 1973; Strupp & Jenkins, 1963; Sushinsky & Wener, 1975; Szasz, 1961; Temerlin, 1968; Wallach & Strupp, 1960; Wright, 1960). This area has also been researched by those sociologists whose attention to this problem focuses upon diagnostic labeling as a function of socio-cultural influences (e.g., Braginsky & Braginsky, 1974; Enke, 1969; Goffman, 1963; Scheff, 1966, 1975; Wolfensberger, 1972). Price (1971) refers to these social scientists as advocates of the "social perspective" of mental illness. The third area of importance has been investigated by social psychologists interested in the fields of person perception, impression formation, and expectancy effects as they relate to the process of clinical diagnosis (e.g., Asch, 1956; Bieri, 1953; Cline & Richards, 1964; Cronbach, 1964; Estes, 1938; Farina, Allen & Saul, 1968; Farina & Ring, 1966; Hastorf, 1970; Orne, 1962; Rosenbaum, 1968; Rosenthal, 1964, 1973; Ryan & Hastorf, 1975).

The practical relevance of research in this area involves the alteration or distortion of one's perception of an individual's otherwise neutral or normal behavior
when a diagnosis is imposed on that individual. This concern is especially crucial for the psychotherapist, professional and paraprofessional, whose aim is the modification of maladaptive behaviors and the strengthening of adaptive ones. Research has shown that certain pretherapy information can have a great deal of bearing on how the client and therapist relate and respond to each other during their initial session (Gustin, 1969). Examples of pretherapy or advance information which many therapists have at their disposal are the sex, the age, the educational level, the race, the ethnic background, the address, and often the diagnosis of their prospective clients. A therapist may receive this information via a report as structured as a complete diagnostic workup to a report as unstructured as a receptionist's casual comments on the appearance of the new client in the waiting room. "Whether the therapist consciously uses this information or not, the fact remains that he/she has it, and it undoubtedly affects his/her attitudes toward the client" (Gustin, 1969, p. 20). Thus, clinicians who are armed with the diagnoses of their yet to be seen clients, may be biased in their relationship with the client in the first therapy session and perhaps in subsequent contacts.

In the case of children, early ascription of diagnostic labels may in Nicholas Hobbs' terms, "generate expectations that often work at cross purposes with the most enlightened efforts to help children" (Trotter, 1975, p. 5).
It is hoped that pertinent research in the realm of clinical diagnosis and judgment will eventually produce practical ideas for alternative categorizing systems, modifications of our contemporary schema, or more insightful and cautious applications of current diagnostic labels.

There have been numerous studies during the last 60 years analyzing the process and results of both child and adult psychotherapy. From these studies, much has been learned about psychopathology and the nature of the patient. Psychologists, these days, have a fair idea about which client variable combinations make a good prognosis. However, what information do we have on ingredients for the "adequate clinician?" Research devoted to elucidating the therapist's part in the process of psychotherapy has been scant (a few notable exceptions: Brenner, 1971; Carkhuff & Truax, 1967; Cicchetti, Ornston, & Towbin, 1968; Goldstein & Shipman, 1961; Spilken & Jacobs, 1968; Whitehorn & Betz, 1954; and Wolpe, 1969). Attempting to add useful information to the sparse literature on this topic, the present study was concerned with characteristics of the clinician--specifically, the possible relationship between the biasing effect of diagnostic labels assigned to children and how such potential biasing may be related to professional ideology and training. It seems possible that clinicians who subscribe to coexisting but rival systems of ideas about the nature and treatment of maladaptive behavior in children could be
differentially influenced by pretreatment information such as a diagnosis.

The present study attempted to explore the variable of suggestion inherent in clinical judgment and how this variable is related to observer characteristics such as professional training, experience, and therapeutic orientation. This investigator was primarily attending to the set a psychiatric label imposes on mental health professional observers of differing theoretical frameworks and undergraduate observers, which influences their perceptions and interpretations of behavior. Since the process of the labeling of handicapped children seems to be of grave concern to many clinicians and administrators, the present research focused on the use of traditional medical diagnoses with "emotionally disturbed" children in a special school setting. This investigator, in previous research, was interested in the alteration of perception, rating, and interpretation of specific behaviors as well as the generation of self-fulfilling prophecies which early imposition of labels could lead to (Saper, Note 6). The present study was also addressing the issue while it attempted to sort out the relationship between the biasing effect of diagnoses and the professional clinician's training and treatment ideology or approach ("traditional" clinicians vs. "behavioral" clinicians vs. undergraduate college students).
Nature and Function of Diagnoses

Scientific classification or taxonomy has been one goal of all physical and natural sciences. Social scientists, and specifically psychologists, have yearned for an organized system of classification of mental health and illness since the genesis of the science itself. Szasz (1961) stated that our preoccupation with identification and classification is fundamental to the need "to order the world around us." The classification schema currently used in the United States, which is one of fifty systems used throughout the world (Zubin, 1967), is one adopted by the American Psychiatric Association in 1968 [modified version, Diagnostic and Statistical Manual of Mental Disorders, second edition (DSM-II), 1971]. It has its basis in Kraepelin's 1883 description and clarification of mental disorders, and was greatly affected by the adoption of the "medical model" of psychopathology.

In the psychological research literature, "diagnosis" as a topic of study is often neglected because clinicians tend to be more interested in results and cures than in categorization (Caveny et al., 1955). The applied
scientist such as the clinical psychologist, the clinical social worker and the psychiatrist "borrows techniques, assumes their infallibility and creates elaborate intellectual rationalizations to justify the diagnoses rather than subject them to careful experimentation" (Caveny et al., 1955, p. 368). During the last twenty years, when such careful experimentation on diagnoses has been undertaken, results have often been disappointing. From this type of research, criticisms leveled against the contemporary diagnostic system are lack of homogeneity, poor validity and low reliability (Hunt, Jones & Nelson, 1962; Edelman, 1969; Koestler, 1975; Sawyer, 1966; Taft, 1955; Yates, 1970; Zigler & Phillips, 1961).

Part of the inadequacy of the system is caused by a lack of consistency in the basis of each classification. Most categories tend to be descriptive of symptom manifestations, while others relate classification to etiology, prognosis, treatment choice, or social conformity. However, it is important to realize that the diagnostic system cannot be designated true or false, but rather useful or not useful in attaining prescribed goals (Zigler & Phillips, 1961). These goals may be description, etiology, or prognosis, and they may differ with the function they serve, i.e., administrative, therapeutic, research, or preventive.

Although some therapists may use diagnostic labels merely to pigeonhole patients, most professionals would
agree that the purpose of diagnosis should be geared primarily towards intervention and secondarily towards heuristic classification. Gough (1971) sees the main goal of psychodiagnosis as the identification of the patient's presenting problems in such a way that the appropriate treatment can be implemented. As an advocate of a "medical model" of psychopathology, Gough feels that if any treatment is to be effective, it must be addressed to the underlying conditions of the disease which should be determined by accurate diagnosis. Levy (1963) also sees the function of psychodiagnosis as more than labeling. "Psychodiagnosis is a descriptive venture, having as its ultimate goal, the provision of a basis for the anticipation of the behavior of the patient under various contingencies" (Levy, 1963, p. 157).

Criticisms against the present use of diagnostic categories other than their reliability and validity, are that current clinical labels tend to be dehumanizing (Winthrop, 1964); they are incomprehensible to anyone outside the field of psychology and psychiatry (Klopfer, 1962); lengthy and clumsy evaluations take crucial time away from psychotherapy (Klopfer, 1962); diagnosticians use arbitrary and ambiguous labels and convince themselves they are scientific (Chein, 1966); diagnostic impressions shift with repeated exposure to the client (Edelman, 1969), and of primary importance to this study, diagnoses may lead the
counselor to erroneous conclusions about the client and retard rather than facilitate the therapeutic process (Berdie, 1950; Gauron & Rawlings, 1973; Sherer, Note 8).

Nicholas Hobbs, in a 1975 report to the Department of Health, Education, and Welfare entitled Issues in the Classification of Exceptional Children, argued for more precise diagnostic practices which are closely monitored in the mental health field. He emphasized that children who have been carelessly categorized and labeled as "different" often become stigmatized for life and are denied the educational and vocational opportunities guaranteed to others. Under the guise of being treated, they are forgotten, neglected and abused. When addressing the notion that diagnoses tend to dehumanize or stigmatize, it is important to realize that most clinicians would argue that this generally is not the goal of psychodiagnosis. Stigmatization or dehumanization are either unfortunate byproducts of diagnosis or the negative effects of diagnoses abused (Gough, 1971). "Like mathematics, diagnosis is susceptible to the deliberate distortion of liars and the unwitting distortions of fools. Diagnosis itself remains guiltless" (Caveny et al., 1955, p. 368).

The diagnostic labeling process also has some ardent supporters. Clarizio and McCoy (1976, p. 112) mentioned the following issues in defense of the current diagnostic labeling practices as they apply to children:
1. Labeling has made it possible to identify major social problems and to marshal vast resources of money, facilities, and talent to attack problems. Without labeling, literally millions of youngsters would never have had any special attention to their needs (Gallagher, 1972).

2. Special programs do not produce cleavage between the special child and peers. The cleavage already exists. A sense of difficulty develops long before any special class placement (Meyers, Sundstrom, & Yoshida, 1974).

3. Although special programs may not be more effective than regular school programs with respect to the three R's, they have been successful in helping to promote employability and self management in the postschool years (Kolstoe, 1972).

4. The notion of a self-fulfilling prophecy has not been substantiated by subsequent research. Nor have the "negative" effects associated with labeling been proven (Kolstoe, 1972).

5. If special programs geared to fit the abilities and needs of youth are charged with being ineffective, how can general education programs provide the necessary services?

Hobbs (1975) has called for definitive research on the labeling phenomenon, especially the investigation of which youth are most susceptible to labeling effects given certain situations and developmental circumstances. The present research attempted to understand a few of the issues involved in the biasing effect of traditional psychiatric labels imposed on children.

With the growth of behavioral approaches in psychotherapy, renewed interest has been stimulated in clinical
assessment and diagnostic evaluation (Goldfried & Kent, 1972). Some psychologists are suggesting that clinicians refer to units smaller than the total personality hoping that the reliability of diagnoses may be greater if the label is based on specified behaviors rather than global personality (Scott, 1968).

Social Perspective

The social perspective of mental illness is probably the most recent to develop, following the psychoanalytic, medical (illness), learning, moral, and humanistic perspectives (Price, 1972). Its major proponents are Goffman, Becker, Scheff, and Sarbin, and they consider social labeling as one of the major causes of deviant behavior. These researchers focus on the diagnosis ascription process, namely, who is labeled as mentally ill, by whom, and under what circumstances. Deviance is not considered "a property inherent in certain forms of behavior; it is a property conferred upon these forms by the audiences which directly or indirectly witness them" (Braginsky & Braginsky, 1974, p. 111). "The deviant is one to whom the label has been successfully applied" (Becker, 1963, p. 18). Social scientists in this field see diagnostic labels not only as "misconceptions of reality but also as misleading and obfuscating . . . it is the labelers rather than the recipients who suffer from poor reality testing and
defective intellectual processes. . . . Labels reveal a
great deal about diagnosticians and the society they serve" (Braginsky & Braginsky, 1974, p. 24). Hobbs (1975), refer­ring specifically to the labeling of children, states that
various child-care systems are controlled by different pro­fessional groups, each of which employs a different category
of exceptionality. How a child gets labeled or "trademarked"
often depends on the professional identity of the labeler.

Scheff's (1966) sociological theory of mental illness
has two basic components, social role and societal reaction.
He assumes that even the most chronic mental illness is in
part a social role and the societal reaction is the most
important determinant of entry into that role. This soci­etal reaction is often organized and activated by a psycho­atric diagnosis since the state is legally empowered to
commit those persons labeled as mentally ill. Scheff (1966,
1975) and Becker (1963) have devised a "labeling theory" of
mental illness. Even though such a theory of deviance is
hypothetical it provides researchers with a useful perspec­tive of abnormal behavior, if only because on many dimen­sions it is diametrically opposed to the "medical model."

Goffman (1963) wrote that the person diagnosed as
mentally ill is "stigmatized" and the stigmatized person
is reduced in the observer's mind from a "whole and ordi­nary" person to a "tainted and discounted" one. However,
mental illness usually consists of symptoms vaguely defined,
and the designation of behaviors as symptomatic of mental illness depends more upon social than upon medical contingencies.

For therapists working with physically or mentally disabled clients, the diagnostic labels tend to perpetrate the "medical" myth of the fragility of all patients. When therapists succumb to this myth and view patients as fragile, therefore easily harmed or damaged, they tend to move too slowly with their interventions and hinder their clients' growth (Gauron & Rawlings, 1973).

With the shift in applied psychology in the last thirty years, from hospitalization towards community mental health, has come a change in the definitions of deviant behavior. Labels assigned to such behavior strongly influence attitudes towards those regarded as "deviant" and the labels tend to activate pre-existing beliefs about the mentally ill which is often to the detriment of the individuals so labeled (Rabkin, 1972). Rabkin does not see the major problem in the mental health field as society's negative evaluation of mental illness, but rather the accompanying rejecting attitude toward the mentally ill and the formerly mentally ill. Cumming and Cumming performed their now classic 1957 study in a small middle class Canadian community, and they found that the general public had a basically negative attitude toward mental illness and it is infeasible to modify this specific attitude without modifying the whole social system.
(Rabkin, 1972). Nunnally found a significant portion of his subjects tended to regard the mentally ill as dangerous, dirty, unpredictable, and worthless (Rabkin, 1972). No recent evidence has been found to support the previous findings (Hollingshead & Redlich, 1958) that attitudes about mental illness are related to educational level or social class (Rabkin, 1972).

The overall negative attitude towards people labeled as mentally ill can also be found towards people labeled as physically disabled. Denise Sherer from New York University who has cerebral palsy recently spoke at a United Cerebral Palsy convention where she poignantly stated that American society treats the person who is physically disabled like a monster. One must realize "that a person's disability is only one aspect of his total being . . . I am not a disabled person, but a person with a disability" (Note 8).

Children are often powerless pawns in the diagnostic labeling process. They are involuntary participants in any evaluation or modification procedure occurring in their school. Undesirable behavior in a child is whatever behavior is regarded and treated as such by his/her teachers. Wickman (1928, p. 4), very early in the literature, called on the clinician to "consider both the child whose behavior is troublesome and the teacher who is distressed or disturbed by the child's conduct." He had teachers make lists
of behavior problems they experienced in their classrooms, and he found a great variation in teacher reports. Apparently there is a large difference between individuals in their observation and accurate labeling of behavior disturbances.

Phillips (1963) cited a study in which a person with symptoms of schizophrenia, but not labeled as such and described as not receiving any help whatsoever, was seen as normal by observers; whereas a normal person, but not so labeled and described as having been in a mental institution, was seen as severely disturbed. Bersoff (1971) proposed that these results may apply to children now found in "special classrooms." If they were called normal and kept in regular classes they would be less rejected by society than if isolated in special classes.

Clarizio and McCoy (1976, p. 111) wrote that the disproportionate numbers of Blacks and Chicanos in special school programs lends support to the charges that these programs are dumping grounds for problem children from minority groups. They cited the following negative aspects of labeling children which are commonly discussed in the literature:

1. Labeling a child prejudices responses of teachers, peers, family, and society which in turn leads to greater behavioral problems.
2. Labeling creates a level of fear which is not warranted by the actual condition--for example, "minimal brain dysfunction."

3. Labeling emphasizes a child's negative characteristics or deficits whereas education and psychological treatment should focus on assets.

4. A label refers to only a fraction of a child's total behavior. A case in point is the "six-hour" retarded child who performs poorly in school but functions adequately at home and in the community.

5. Labels on children lead to a neglect of individual differences and towards a "two-box" theory of education.

6. Labels can create a change in others' expectations for a child which often creates a "self fulfilling prophecy."


8. Labeling often leads to exclusion from social systems rather than remediation. Children become warehoused instead of treated.

9. Children who are labeled "emotionally disturbed" through a diagnostic procedure which may be invalid or unreliable, may be unnecessarily exposed to inappropriate and sometimes dangerous peer models.

10. The labeling of a child implies that the problem is within the child, often leaving the influence of the family, school, and community inattended to.

11. Placing a child within a diagnostic category often makes it difficult for a child to move out of it in spite of a change in his/her condition.

Two important and relevant investigations using social perspective hypotheses are Temerlin's (1968)
research on "suggestion effects in psychiatric diagnosis" and Rosenhan's (1973) study involving the experiences of eight "normal" pseudo-patients who gained admission to twelve different mental institutions.

Rosenhan had eight pseudo-patients admitted to the facilities by giving their veridical social histories and saying that they were having auditory hallucinations. Eleven of the subjects were diagnosed as schizophrenic and one was diagnosed as manic depressive. Rosenhan stated that the important issue was the diagnostic leap made between a single presenting symptom and the diagnosis of mental illness. He suggested that the description, "hallucinating," was all that should have been warranted by the admitting physicians and by our current state of knowledge. Once the person was designated abnormal, the perception by hospital staff of his/her other behaviors and characteristics was colored by that label. It took two weeks before most of the pseudo-patients, who behaved as their normal selves on the ward, were discharged. They were discharged with the diagnosis of "schizophrenia in remission." The ward aides often recorded the behavior of the subjects as abnormal (such as compulsive note taking) which outside of the mental hospital would have been considered a normal activity for researchers. This finding might suggest that the appropriateness of a behavior is not independent of its setting.
Rosenhan did a corollary study in which he told staffs of eight hospitals that at least one pseudo-patient would try to get admitted to their hospitals within the next three months and they were to rate from one to ten the probability that each admission was a fake. No pseudo-patients were actually involved and a significant number of admitting physicians rated actual patients as fakers.

Rosenhan received wide public acclaim over his studies as well as numerous rebuttals from scientists in the field. His data and results seemed to be more appealing to the lay population than to professional researchers. Several of Rosenhan's critics presented reasonable faults with his research. Fleishman (1973) suggested that the pseudo-patients did fake histories and therefore were diagnosed correctly on the basis of those histories. Most doctors do not expect voluntary admissions to be liars. Ostow (1973) reported that mental illness can be easily simulated and note-taking compulsivity is common among patients in hospitals. If a doctor refuses to admit such a person into a hospital, and that person later commits suicide or homicide, the doctor could be in legal trouble. Lieberman (1973) stated that Rosenhan's study actually proved that competent judges cannot distinguish the insane from the sane feigning insanity, when judges are aware of no reasons to feign insanity. Hunter (1973) wrote that the pseudo-patients on the ward did not really behave
normally, since a normal person would have gone to a nurse's station and said, "I am a normal person who acted crazy to get in here and would now like to get out."

Rosenhan was inspired to undertake his investigation by some research completed by Langer and Abelson (1974). They video-taped an interview in which discussions were focused on a client's job history and difficulties. Then two groups of observers, one consisting of trained psychodynamic psychologists and the other consisting of behavioral therapists, were asked to rate the degree of adjustment of the client. Half the group were told they were watching a psychiatric interview and the other half that they were watching a job interview. It was hypothesized that the therapeutic orientations of clinicians would influence the effect the labels had on their judgments. It was thought that therapists who were behaviorally oriented would be quite skeptical about the utility of diagnostic categories and labels. Those psychodynamic clinicians who thought they were watching the job interview rated the subject as better adjusted than those who thought they were rating a patient in a psychiatric interview. The effect of the label was non-significant between the groups of behavioral clinicians.

Temerlin (1968) had groups of psychiatrists, clinical psychologists, and graduate students in clinical psychology diagnose a tape of an actual clinical interview of
a "normal, healthy person" played by an actor. Just prior to listening to the recording they heard a professional person of high prestige say that the individual to be diagnosed was "a very interesting man because he looked neurotic but actually was quite psychotic." The credible source for the psychologists and graduate students was a well-known psychologist with many honors; and the credible sources for the psychiatrists were two board certified psychiatrists and one psychoanalyst. After listening to the patient the judges indicated their diagnosis from among ten psychoses and ten miscellaneous personality types, one of which was "normal." Each judge also wrote a brief clinical report of the patient to indicate the behavioral basis for his or her diagnosis. They were instructed to avoid inferences.

None of the control subjects for whom no diagnostic label was presented diagnosed psychosis while diagnoses of psychosis were made in the experimental groups by 60 percent of the psychiatrists, 28 percent of the psychologists and 11 percent of the graduate students. In their clinical reports most subjects either mixed inferences and observations or reported inferences exclusively. Only the few subjects who diagnosed health reported straight observations or behavioral data. Temerlin concluded that suggestion effects contribute to the unreliability of psychiatric diagnoses. Apparently diagnostic labeling varies with
personal values, training, and perceptual consistencies of the individual diagnostician.

In three replications and extensions of Temerlin's work, Sushinsky and Wener (1975) reported the following results: Labeling bias was not found in undergraduates when there was no prestige figure presenting the diagnosis; however, labeling bias was produced in undergraduates by manipulating the "relevance" of the prestige figure, and, labeling bias was demonstrated in mental health workers in a psychiatric hospital. In one of these experiments these researchers supplemented Temerlin's design by utilizing a taped interview of an actual psychiatric patient along with Temerlin's "normal" interview in order to assess the rater's ability to discern and rate according to the actual information being transmitted (Sushinsky & Wener, 1975). They found that the labeling bias effect was a general phenomenon, but also found that undergraduates could discriminate an audio tape recording of a psychiatric patient from a tape of an actor playing a normal person.

The Temerlin study and its replications as well as the Langer and Abelson study were well controlled laboratory experiments. Rosenhan's study, on the other hand, was a field study lacking in experimental controls and sophisticated methodology. Each type of research has its obvious benefits and limitations in exploring the social perspective field.
Person Perception and Expectancies

Social psychologists have been interested in the area of person perception since the early work of Estes in 1938. He studied how accurately observers judge the personality of subjects from their expressive, non-verbal behavior. He used two-minute film clips of actors who walked into a room, removed coats, played Black Jack, built houses of cards, etc. The situation provided an opportunity for the actors to demonstrate a variety of expressive movements which were representative of their behavior in real life situations. Judges then rated the actors on personality dimensions or selected appropriate personality descriptions for each actor from several possibilities. Estes found that judges varied widely in their ability to match the behavior of actors with their personality sketches. Variance in accuracy was associated with the characteristics of the judge, the characteristics of the subject, and the aspects of personality being judged. Judges with strong interests in the arts were more accurate than those with interests in the sciences or philosophy. Adult judges, in general, were more accurate than student judges. There was more accuracy when judges were asked to make global judgments by matching rather than rating scales; and there was greater accuracy when subjects tended to be expressive rather than introverted.
Cline and Richards (1964) suggested that there is a general ability to perceive others accurately. This ability consists of at least two independent parts, sensitivity to the generalized other, and interpersonal sensitivity, (Bronfenbrenner's terminology), or in Cronbach's (1964) terminology, stereotype accuracy, and differential accuracy. Cline also employed film in his research. He had color film interviews of ten different people. Judges were given the task to postdict possible real life behavior of the person seen in the film. (For example, "When X gets angry, he usually ________.") These items were tailored to each film and responses were rated by clinicians who had earlier tested the film subjects, interviewed them, and acquired background material on them. Cline's findings lended support to Cline and Richards' (1964) two component theory of person perception.

Allport (1955) has written that the ability to accurately judge behavior is like an artistic ability which is neither entirely general, nor entirely specific. Hastorf (1970) suggested that some dimensions of rating seem to lend themselves to accuracy more than others.

Perhaps a two-minute silent film clip (Estes, 1938) provides too little information to obtain any accuracy in judgment; yet too much information could be confusing (e.g., sound films, Cline, 1964; live observations, Wickman, 1948).
In one experiment, Cline (1964) showed films of four highly structured interviews with college males to groups of judges. During each 11-minute interview, as much objective and emotional information as possible was elicited from the interviewee. He had his judges fill out three measures involving prediction and postdiction of the subjects' behavior and responses. Cline also questioned whether the judges would have been as accurate in their responses had they not seen the films but instead had responded to the same three measures according to their stereotype of a typical college male. He was primarily interested in Cronbach's component of "stereotype accuracy." Cline ran a control group of 57 undergraduates who completed the same prediction and postdiction measures by guessing what they felt a typical college male would be like. Cline obtained significant results (p < .001) with two of his three measures favoring those who had seen the films. This evidence suggested that the judges watching the films were making accurate predictions or ratings on the basis of differential analysis and a real evaluation of the personalities of the film subjects, rather than from a crude internalized stereotype of what college males were like. On the average, the group of judges who were professional clinicians proved most accurate, followed by a group of judges who were nurses, a group who were college students, and a group who were church members and engineering trainees.
Rosenthal and Orne dealt with person perception and suggestion effects from the angle of "demand characteristics" of the experimental, educational, or therapeutic situation. Orne (1962, p. 77) defined demand characteristics as the "totality of cues which convey the experimental hypothesis to the subject and which become significant determinants of subjects' behavior." Rosenthal examined an aspect of this phenomenon in his study of the effects of the experimenter on the results of psychological research. He found that observer bias tended to produce results consistently much lower or higher than a true or criterion value. Observer bias is related to characteristics of the observer or the observation situation or both (Rosenthal, 1954). In one of his earlier studies, Rosenthal had students rate photographs of people on a scale which ran from "experienced success" to "experienced failure." The subjects were told that the experimenter wished to see whether they could replicate well-established experimental findings, as students in physics labs are expected to do. Depending in which direction the experimenter said the findings were expected to go, ratings by the students were consistently and significantly in that direction on the scale.

In later studies utilizing a classroom setting, Rosenthal found that teachers who expected certain students to perform better than the rest of the class because of
information given them from a variety of sources (psychologists, tests, etc.), create a warmer social emotional climate around those students, give them a larger amount and more difficult material than given to the others, give more feedback to them, and give them more opportunity to respond to questions. Thus a child who is expected to perform well, generally will. Also of importance is the fact that if a child who is not expected to perform well, does so, his teacher will often look upon his behavior and personality as undesirable. This is especially true of children in low ability classrooms (Rosenthal, 1973).

Rosenthal and Orne were primarily interested in the expectancy effect as it related to experimenters, teachers, and college student subjects. Wright (1960), Goldstein and Shipman (1961), and Gustin (1969) have researched the therapist's expectations of the patient in psychotherapy. Goldstein wrote that therapists cannot hope to understand their clients' states of mind or their behavior unless they consider their own expectations about themselves and those with whom they interact.

Helmut Enke (1969), a German psychologist, wrote that therapists are as subjective as everyone else and this influences their diagnoses and the modes of treatment their patients receive. The psychotherapist is a member of an elite minority group which projects universal pretensions, roles, and images.
Wallach and Strupp (1960) found that a therapist's positive attitude toward a patient was associated with a favorable diagnosis and prognosis. This positive attitude also contributed toward the therapist's estimate of greater patient ego strength, insight, and greater ease of anticipated empathy with the patient. In Strupp's later research (Strupp & Jenkins, 1963), he had professional clinicians watch a film of a staged initial interview. At various times during the showing the film was stopped and the audience had to decide what they as therapists would do next.

Gustin (1969) using Strupp's 16mm sound film designed an investigation to determine whether therapists were subject to expectancy effects, on what dimensions biasing occurs, if in fact it does occur, and the way the phenomenon affects the therapist-patient interaction. Gustin's subjects were advanced graduate students in clinical and counseling psychology. The study was concerned with the effect case history information had on the therapists' perceptions of, and attitudes towards the filmed patient. All subjects viewed the same film of the patient/actor in a psychiatric interview. The case history material presented to the subjects varied on two different dimensions--the diagnosis assigned to the patient ("psychopath," "neurotic," or not stated), and his purported motivation for therapy (high motivation, low motivation, or not stated). Measures used were Strupp's Prognosis-Evaluation Scale, a
scale to judge therapists' commitment to the patient, and open-ended responses to the patient/actor at points when the film was interrupted. From these open-ended responses to the films, judges measured the therapists' warmth and empathy on scales by Carkhuff and Truax (1967).

Gustin found that therapists seemed to relate to the filmed patient diagnosed "neurotic" with more empathy, more warmth, and were more attracted to him, than to the same filmed patient diagnosed "psychopath." Commitment to the patient and prognosis were not significantly affected by the diagnoses, but were affected by the perceived state of motivation. Gustin's study clearly demonstrated that therapists are subject to expectancy effects and that this bias affects their behavior toward clients, how they feel about them, and how they perceive their behavior. Of most interest to this investigator, Gustin found that the diagnostic label was extremely effective as a biasing factor.

The influence of perceived mental illness on interpersonal relations has also been studied by social psychologists (e.g., Farina, Allen, & Saul, 1968; Farina & Ring, 1966). They found that the belief that a person is mentally ill strongly influenced the perception of that individual, even though the "ill" person's behavior in no way justified these perceptions. They concluded that when a person is "stigmatized" (they used the labels "mentally ill" and
"homosexual" in various studies), he/she is not only evaluated less favorably, but also, other people behave differently and generally less favorably toward him/her. In Farina's (1966) earlier study, subjects read fake biographies of their partners prior to doing a task. There were two different biographies which subjects read depending on which group they were randomly assigned to. One biography read: "I have certain problems in adjustment... I was placed in a mental institution when I had a kind of nervous breakdown." The other biography read: "I tend to think of myself as relatively normal..." (p. 20). The subjects in this study were young and well educated, characteristics once thought to be associated with favorable attitudes toward the mentally ill; nevertheless, results showed subjects to react more unfavorably to "ill" partners than to normal partners. It is apparent that under certain circumstances what a person supposedly reveals about himself/herself significantly influences the perception of his/her behavior by an observer even though the actual behavior does not justify that perception. One argument against the conclusions drawn from this study is that the type of encounter was so brief and superficial that perhaps what was measured were people's stereotypes about the mentally ill. However, this initial impression would probably reduce the chance for further interaction which keeps the mentally ill and stigmatized person a "prisoner of his own reputation" (Farina & Ring, 1966).
In person perception terms relevant to the present investigation, Ryan and Hastorf (1975) conceptualized the results of studies by psychologists such as Farina and Ring in the following way: Observers are willing to go beyond the behavior they actually observe when they formulate theories about what other people are like. Inferences made by these observers will usually be consistent with and related to certain critical information which they feel they already possess (e.g., diagnostic label, motivational disposition, past history or performance).

... Individuals when they observe others see much more than simply physical acts. They select the information to which they will attend, they construct categories into which they sort this information, and they place an interpretation upon the resultant "events" (p. 3).

The Nature and Behavioral Observation and Rating

A typical explicit or implicit sequence that a clinician follows before he/she begins treatment is to carefully observe the client's behaviors and then to rate those behaviors on maladaptive-adaptive scales. The clinician ordinarily follows these steps before attaching a label to the client. Blumberg (1971) devised a training program in behavior observation to be used at the United States Army Medical Field Service School. He divided observations into three categories: (1) visual--facial expressions, body posture, and behavioral gestures; (2) auditory--rate, volume,
and tone of voice plus vocal gestures; (3) tactile cues—these are usually minor. Blumberg emphasized that the presence of one sign of behavioral abnormality, even when it is quite clear, does not warrant classification of the client's subjective state. Validity of a diagnosis is achieved through a number of signs or cues. An example of this concept would be that the observation of someone smoking might persuade the observer to deduce anxiety, though a deduction of this kind may be premature. In Blumberg's training programs he illuminated features which keep clinicians from objectively observing what is actually present: (1) preconceived ideas of the observer, including set, biases, and prejudices; (2) personal needs of the observer and "self validating phenomenon"; (3) situational factors such as economic or social gains that can motivate clients to behave in ways inconsistent with their feelings outside of the observational setting. The primary concern in the present research was with feature number one (specifically the set of diagnosis).

During observation of people's behavior, various types of judging instruments can be used by the clinician. The following is a partial list of such instruments used in person perception research:

1. trait-rating procedures such as adjective checklists, semantic differentials, Likert-type rating scales;
2. postdicting real life behavior, usually true or false or multiple choice questionnaires;

3. postdicting responses to specific objective test items, for example, Minnesota Multiphasic Personality Inventory, Strong Vocational Interest Blank, etc.;

4. postdicting scores on achievement tests;

5. postdicting theoretical constructs (psychiatric diagnoses);

6. writing global descriptions of the person being judged;

7. matching person being judged to personality description;

8. ranking procedures based on the degree of a trait or characteristic present;

9. forced choice tests, for example, the judge predicts which of two statements the subjects would agree or disagree with;

10. Q-sort technique;

11. giving an open-ended therapeutic response to the judged person's statements or behavior;

12. any combination of the above (Cline, 1964).

Most researchers in this area, however, develop their own judging tests in the absence of valid procedures. Whereas Blumberg listed three features which prevent clinicians
from making objective observations, Cline followed this lack of objectivity to the next step and offered ten causes of errors in judgment. He focused on the kinds of biases and response sets which influence rating:

1. social desirability or the tendency to predict the most socially desirable response in making predictions and judgments of others;

2. similarity of the judge to the subject being rated;

3. acquiescence set;

4. employing an undifferentiated stereotype to predict the behavior of the person being observed;

5. personal reaction of the judges such as liking or disliking the individual which can produce a "halo" effect in rating and judging;

6. making use of an implicit personality theory: for example, assuming there is an invariant relationship between trait "A" observed in the subject and traits "B," "C," and "D" not observed;

7. central tendency response set;

8. the assumption of the judge of similarity to the subject, creating projection;

9. the assumption of the judge of dissimilarity to the subject, creating projection "in reverse";
10. **semantic ambiguities** which cause the judge to interpret a trait name in the rating instrument to mean something other than it was intended to mean in its development and use.

Of relevance to the present study is a combination or interaction of the above items which may be ingredients of a clinician's theoretical framework of abnormal behavior. The most important components of such a framework are personality variables and professional training experiences. These variables may contribute to the accuracy or inaccuracy of a clinician's observations, judgments, and behavior ratings. Such clinician variables, and specifically theoretical ideology, which affect these processes will be discussed in detail under the next subheading.

**Clinician Variables Which Affect Behavioral Observation, Rating, and the Therapeutic Relationship**

Thus far in reviewing the relevant literature, diagnostic labeling has been discussed as a functional tool of clinicians, as a sociological phenomenon, and as an important factor in person perception and expectancies. Several important studies in each of these three areas have been explored. Most of these investigations dealt with the "demand characteristics" of the labeled subject/client or situation. It is now appropriate to look at the personal characteristics that the rater, observer,
or therapist brings to bear when responding to the labeled person. Allport (1955) and Cline (1964), as mentioned earlier, feel that there is a "general ability" to perceive others accurately. Theoretically, therapists may or may not have this ability. It seems probable that there are personal variables which might compose such a capability and which affect a clinician's perception and judgment in a therapeutic situation. If these characteristics affect the therapists' perceptions of their clients, one could hypothesize that therapists with varied combinations of personal characteristics are differentially influenced by pretherapy diagnostic labels.

Strupp (1958, p. 219) wrote that "the totality of the therapist's personality: age, sex, experience, maturity, attitudes, etc. are subtly intertwined with his therapeutic techniques and the theoretical framework he brings to bear upon his therapeutic operations." Some recent investigations to be discussed next have been completed on a few such therapist personality variables--empathy, experience, authoritarianism, objectivity, dependability, sincerity, directiveness, respect, and warmth. However, only a handful of researchers have looked into a therapist's "theoretical framework" (Strupp, 1958) or training as it influences his/her perceptions of the labeled client.

Carkhuff and Truax (1967) found that all theories of therapy stressed a therapist's ability to be accurately
empathic, warm, genuine, and not possessive. Spilken and Jacobs (1968) found that therapists valued empathy, respect, and interest more, and objectivity, dependability, sincerity, sureness, and directiveness less than non-therapists. Inexperienced psychologists tended to value empathy more than experienced psychologists and psychiatrists. Dubnicki (Note 2) found that there was a positive relationship between the therapist personality trait of empathy and the therapist's perceived prognosis for the client. The relationship between empathy and perceived degree of disturbance was a negative one, implying that the more empathic the therapist, the less pathology will he/she perceive in the client.

Cicchetti, Ornston and Towbin (1968) found that novice therapists used more questions and fewer words in their responses to filmed clients than experienced therapists. However, they found that differences in responses between experienced and novice therapists eventually disappeared by the third contact with the filmed client. Wolpe (1969) found that experienced therapists were more effective in decreasing clients' anxiety than were novices. Strupp (1958) found that experienced therapists tended to be "warmer" than novices. Brenner (1971) found that experienced therapists were better able to assess their own accuracy in empathizing with clients than were inexperienced therapists.
A somewhat different finding came out of 1965 re-
search by Breiter, Golann, and Margoon (Note 1). They
found that housewives with two years of training did bet-
ter at empathizing with inpatient clients than did hospital
volunteers. The housewife group appeared more similar to
the experienced therapist group. The difference between
groups was not in the amount of contact with the patient,
but in the amount and type of training the groups had.

Another way of describing a therapist's type of
training is to use Strupp's (1958) notion of the therapist's
"theoretical framework." Vardy (1971) refers to this con-
cept as the therapist's "ideology." He sees the clinician's
eventual commitment to a specific clear-cut ideology as a
multidetermined process. Developing personality, teachers,
social value systems, contemporary social milieu, and ex-
periences, all contribute to the adoption of a specific
professional ideology by the psychotherapist. The word
ideology has been used to designate an encompassing system
of ideas. Price (1972) refers to this concept as a "per-
spective" of abnormal behavior. Perspectives are metaphor-
ical, emphasizing the perceptual and organizational aspects
of a thought process. "Because adherents of differing
views of abnormal behavior seem to experience the same
events in radically different ways and because they tend
to see the behaviors and events in question in terms of
their own metaphor. . . ." Price categorized them in
different perspectives (illness, psychoanalytic, moral, learning, humanistic, social). "The same set of puzzling behaviors viewed from two different perspectives may have little or no overlap in terms of the events which are considered by each perspective to be relevant" (Price, 1972, pp. 15-16).

The perspective of abnormal behavior a clinician subscribes to has other implications besides how he/she views, selects, and interprets a client's relevant behavior. Vardy (1971, p. 547) wrote:

The adoption of a specific professional ideology by the psychotherapist . . . has strong implications in terms of its functions as a symbol of his social belonging and of his group membership. The adherence to a certain professional sub-ideology tends to designate and define the position of its holder on a spectrum of professional issues and also indicates his place among the professional factions.

Pasamanick, Dinity, and Lefton (1959) studied two psychiatrists on the same ward of a mental hospital over a two-year period. During this time they gave diagnoses to the same patients. One psychiatrist diagnoses "schizophrenia" in 22% of the patients and the other diagnosed "schizophrenia" in 67% of the patients. The researchers offered their data as statistical affirmation that:

Clinicians may be so committed to a particular psychiatric school of thought, that the patients' diagnoses and treatment are largely predetermined. Clinicians, as indicated by these data, may be selectively
perceiving and emphasizing only those characteristics and attributes of their patients which are relevant to their own preconceived system of thought. As a consequence, they may be overlooking other patient characteristics which would be considered by colleagues who are otherwise committed. This makes it possible for one psychiatrist to diagnose nearly all of his patients as schizophrenic while an equally competent clinician diagnoses a comparable group of patients as psychoneurotic (Pasamanick et al., 1959, p. 131).

In his research utilizing 30 psychiatry residents, Vardy (1971) divided clinician "ideology" into three categories: somatotherapeutic; psychotherapeutic; and sociotherapeutic. Armor and Klerman (1968) had originally designed those categories for their research.

Other investigators have classified orientations differently from Price, Armor and Klerman, and Vardy. Berzins et al. (1971) divided professional orientations into insight, relationship, and action. McNair and Lorr (1964) hypothesized three dimensions: psychoanalytic versus nonanalytic; impersonal versus personal, directive versus nondirective. Hollingshead and Redlich (1958) divided psychiatrists into directive-organic and analytic-psychological. Whitehorn and Betz (1954) suggested that therapists' orientations can be differentiated along a bi-polar (A-B) dimension. The "A-therapist" is primarily concerned with personality oriented goals in treatment, while the "B-therapist" is problem centered and more concerned with symptom
reduction than with alteration of basic personality structure or dynamics. Gilbert and Levinson (1956) found two continuums of thought useful for categorizing a professional's ideology, the custodial-humanistic continuum and the authoritarian-egalitarian continuum. Langer and Abelson (1974) distinguished therapists as behavior therapists or traditional therapists.

It is generally felt that the professional ideologies or theoretical orientations that clinicians have are products of the interaction between their personalities and their training. Empirically, only a small number of individuals identify forcefully and totally with one idea system. The majority of mental health professionals are eclectic to varying degrees. However, using many different instruments, the researchers mentioned above have managed to arrive at methods to categorize clinicians by their perspective or ideology of abnormal behavior and therapy. Several investigators utilize a self-report such as, "What label would you give to the kind of psychotherapy you practice?" (Armor & Klerman, 1968; Berzins et al., 1971; Langer & Abelson, 1974; Weiss, 1973). Other self-reports such as "List the three authors (or books) who have shaped your present therapeutic approach," have also been employed (Berzins et al., 1971). Looking more closely at the personality variables involved in orientation, Whitehorn and Betz (1954) devised a scale from items on the Strong Vocational
Interest Blank and the California Personality Inventory to assess therapists' theoretical orientation on the "A-B" dimension. McNair and Lorr (1964) based their AID scale on the Therapist Orientation Questionnaire (TOQ) devised by Fey in 1958 and revised by Sundland and Barker in 1962. On this measure, a clinician indicates his/her agreement or disagreement on an eight point scale with statements of how therapists should conduct therapy and which therapeutic techniques should be used during interviews.

Vardy (1971) asked two questions of his clinician subjects to assess their ideology. First, what are the best modes of effecting change in psychotherapy by order of importance--catharsis, insight, learning more adaptive behaviors, corrective emotional experience through contact with the therapist, or advice by the therapist? Second, what are the five most desirable characteristics of a psychotherapist? Langer and Abelson (1974) asked their clinician subjects how strongly they agreed or disagreed with the following four statements which touch issues of difference between schools of psychotherapy:

1. If you have cured the symptom you have usually solved the problem.
2. The examination of childhood experience is essential to effective psychotherapy.
3. The use of official APA diagnostic nomenclature for psychiatric disorders is helpful to both patient and clinician.
4. Most people need some kind of psychotherapeutic help (p. 6).
The results of investigations comparing therapists of different ideologies or therapeutic orientations are varied. Much of the variation is caused by different classification schemata, diverse measures, and assorted samples and settings over the last twenty years. Berzins et al. (1971) found that one third of the psychiatrists they studied endorsed an analytic-impersonal-directive model of therapy. One third of the psychologists endorsed a nonanalytic-personal-nondirective model of therapy. The remaining psychiatrists and psychologists were scattered evenly. The social worker sample was split between the two models described above. In their study female clinicians tended to be more impersonal and directive than male clinicians. Experience had no relationship to the three dimensions, yet involvement in personal therapy was related. Psychotherapists who had been in therapy themselves tended to be more analytic than those without that experience.

In an article by Weiss (1973), which he admits may be biased by his own analytic orientation, he related a sketchy, less than experimental, study of 40 therapists in training in which he found: (1) analytic students were more interested in therapy and diagnosis than behavioral students; (2) analytic students made more global statements about clients' feelings than behavioral students; (3) behavioral students were far less interpersonally sensitive than
analytic students; (4) analytic students brought to bear less intellectual acuity where humanistic variables were concerned than behavioral students; (5) behavioral students tended to ignore "unobservables" like personal values and attitudes of the therapist; (6) behavioral students were more oriented towards ideas and overtly observable events than analytic students; (8) both groups were relatively anxious but the analytic students tended to be aware of it, whereas the behavioral students were not. In conclusion, Weiss stated that it was difficult to extrapolate personality factors from the demand characteristics or orientation slant of clinical training programs.

Probably of greater relevance to the present investigation into labeling effects is the Langer and Abelson (1974) study. These researchers found that behavior therapists were more immune to the biasing effects of the label, patient versus job applicant, than traditional therapists. The 21 behavior therapists were from S.U.N.Y. at Stony Brook which has a totally behaviorally oriented clinical psychology training program. The 19 analytic therapists were from N.Y.U. and Yale clinical psychology programs, both of which make no mention of behavior therapy in their program description, and both of which have the objective of familiarizing their students with the theories and practice of dynamic psychotherapy (see pp. 19, 40 for more information on this study). Langer and Abelson hypothesized
that the behaviorists were less biased by the taped subject's label than the traditionalists because they tend to focus on manifest behaviors and inattend to background information such as a diagnosis which they view as irrelevant in making a behavioral assessment of a client.

Brenner (1971) also found that therapists were more accurate in their clinical judgments when they focused on information, behaviors, and cues that the clients were aware of than when they focused on subtle cues, nuances, dynamic formulations, etc. This finding suggests that a behavioral approach to psychotherapy may lead to more accurate clinical judgments than global personality or analytic approaches. Mischel (1968) has also pointed out that in terms of assessment it seems more profitable to discuss behaviors, the stimuli that provoke them, and their correlates, than to discuss global personality traits.

In attempting to integrate much of the previous research on the clinician's therapeutic orientation, this investigator employed a measure to sort therapists into two large ideological categories. One category which was classified as "traditional" is a loose combination of Price's (1972) psychoanalytic and illness perspectives. The other category which was classified as "behavioral" is a loose combination of his learning and social perspectives.
Methodological Issues

The methodologies employed in investigations of clinical judgment and labeling effects are varied. It is helpful to understand some of the methodological problems other researchers in the area have encountered in order to see more clearly the rationale for the methodology used in the present study. All "person perception" and "social perspective" research into the influence of labels includes some type of stimuli to be evaluated, some form of behavior evaluation and set measurement technique, and some diagnostic set induction technique. All research into clinician variables and therapeutic perspective necessarily include some method for sorting therapists by orientation.

As briefly discussed previously, researchers have employed stimuli ranging from still photographs (Rosenthal, 1964) to live stooges (Rosenhan, 1973). The question arises in such investigations as to how representative of real life the stimuli are, and further, how much information the stimuli should emit so that judgments of the behavior witnessed will be accurate. Researchers have to decide how long the stimulus presentation should be and whether it should be in McLuhan's (1967) terms, a "hot" or a "cold" medium. In other words, how much information will the observer have to project of him/herself onto the stimulus in order to evaluate it. One would think that less projection is done when a live person is the stimulus.
than when a still photograph is the stimulus. The present study has adopted a compromise approach to the "amount of stimuli and information issue" found in past investigations, by using silent, color films in which the child subjects rapidly adapted to the cinematographic situation. Since Cline (1964) found that the maximum time judges could sit through films of subjects and accurately take tests was two hours, the present investigator used two films which were each eight minutes long and a 45-minute testing period.

The subjects who make up the stimuli utilized for observers to evaluate is another crucial methodological issue. When films or recordings are employed, they are generally of an actor asked to behave "normally" or read a script of an interview situation; or they are of a "normal" person in a specified situation. What usually varies in such studies are the labels attached to the actor or filmed subject, or the label of the filmed situation. When a live stimulus is used in such investigations, the subject is generally an actor or "normal" person instructed to display phony symptoms or divulge distorted "presenting complaints" for evaluators; or evaluators are given an inaccurate diagnostic label or "expectancy" for the subject. The current study employs two films--one of a normal subject and one of a disturbed subject. This methodology permits comparison of a filmed
subject with him/herself; the only factor changing is the label ascribed to him/her. (Another innovation in this study is that the actual diagnostic label of the disturbed subject determined by the city Board of Education is one of the labels imposed on both filmed subjects.) Such comparisons allowed the investigator to draw conclusions about the effect the imposition of diagnostic labels had on both normal and disturbed children. The present investigation employed films involving children, which is not true of most other investigations in this area of research. Most of the studies on labeling bias utilized filmed, taped, or live adults as stimuli. As discussed previously, children are repeatedly caught at the powerless end of the professionals' ratings, judgments, and diagnostic processes. Probably more often than adults, their behavior is interpreted and labeled by clinicians and they find themselves without the opportunity or capability to appeal or verbally counter decisions made about their future. Thus, the films in this investigation may be unique not only because they feature an actually normal person and an actually disturbed person compared against themselves, but also because these people are children, and children have been long neglected in the research literature on "labeling effects."

The types of instruments used by investigators in the past to evaluate observers' perception of the stimulus
employed were listed on page 31. They include projective techniques, trait-rating procedures, clinical descriptions, etc. Attempts to devise a measure which allows the observer to evaluate the stimulus subject, and can be used to measure the effect of the set which the investigator has induced (e.g., diagnostic label), as well as one which is easily quantifiable, has not been an entirely successful task in the past. A study such as the present one must employ measures which can detect observers' expectations for the filmed subjects (Rosenthal, 1954), detect observers' global perception of the subjects' emotional adjustment (Gustin, 1969; Langer & Abelson, 1974; Sushinsky & Wener, 1975; Temerlin, 1968), plus detect differences in observers' perceptions of specific behaviors—whether they felt behaviors occurred at all, and whether or not the behaviors are interpreted as being normal (Rosenhan, 1973; Wickman, 1928). For this reason the present study adopted three instruments: a global trait rating scale (semantic differential), a problem checklist which would pick up expectancy effects (Peterson Problem Checklist), and a behavior description test closely linked to the actual filmed behavior of the subjects (designed by the investigator).

Social perspective researchers in this area also had to devise a method for inducing a psychological set in observers or evaluators of their stimuli. The challenge
to these investigators has been to employ a technique which
is effective, requires a minimal amount of deception, and
is ethical. The set induced by most researchers investigat­ing labeling effects has generally been a diagnostic
category. Rosenhan (1973) imposed the diagnostic label on
his stooges by their own self report. Temerlin (1968) used
the diagnostic statement about the subject from a "credible
source." Phillips (1963) used written character descrip­tions to induce the diagnostic set. Gustin (1969) put a
statement in a hypothetical staff report regarding the
filmed subject's diagnostic label and motivation for psy­
chotherapy. Langer and Abelson (1974) chose to label the
stimulus situation rather than the actor within it to in­
duce the set in their observers. All these studies em­
ployed deception to a certain degree as does the present
investigation. Rosenhan's study probably involved more
distortion, faked information and interference in the
delivery of actual mental health services than did the
others—reasons that some of his critics suggest that
such research is unethical (e.g., Fleishman, 1973; Ostow,
1973). Through the use of films, the filmed subjects'
actual diagnostic labels, and other procedures used in
the present study (see Methods section) this investigator
hoped to keep deception to a minimum and not infringe
upon the rights or freedom of the observers or the chil­
dren in the films.
The few existing investigations concerned with the theoretical orientation of clinicians and how it affects processes such as therapy or the perception and interpretation of behavior have generally been poorly controlled, biased, and overall, experimentally inadequate. Often researchers have equated the reputed ideological orientation of a training facility or program with the theoretical perspective of the individual therapist (e.g., Langer & Abelson, 1974). Other researchers have been more interested in comparing and contrasting professions rather than therapists' personal theoretical frameworks (Cicchetti, et al., 1968; Spilken & Jacobs, 1968; Temerlin, 1968). However in the last 10 years, the traditional therapeutic professions of psychiatry, psychology, and social work have begun to encompass practitioners with varied and divergent ideas about abnormal behavior and psychotherapy. Currently, much overlap can be found among the theoretical premises of the three professions. It is no longer accurate to conceive of each profession as having its own unique school of thought on mental disturbance. In fact, Armor and Klerman, (1968) in their research found little relationship between social or professional background and treatment orientation.

In the present study the investigator draws on several of the categorizing techniques discussed previously. The Clinician Questionnaire (Appendix B) designed to sort observers/therapists focuses on the individual
clinician's self-report as to his/her therapeutic orientation, goals for the client in therapy, activities engaged in during treatment sessions, preferences for the type and method of mental health intervention, and his/her utilization of psychiatric diagnoses. Although not limiting the investigation to therapists' professional titles or the reputed ideological orientation of the facility they are operating in or were trained in, these two pieces of data were also collected and considered in the categorization of the professional sample (see Materials section for further information on the Clinician Questionnaire).

Pilot Study

In light of the methodological issues discussed and the research completed in the three overlapping content areas (clinical diagnosis, social perspective, person perception) feeding into the exploration of "labeling effects," the present investigator tested the following hypothesis in a 1975 pilot study (Saper, Note 6):

Undergraduate observers who view both a normal child and a disturbed child on film, perceive the children and the children's behaviors as more "abnormal" if told that they have been diagnosed "severely emotionally disturbed . . . pre-psychotic symbiotic ties, mild mental retardation, and epilepsy" than if told that they are "normal."

The observers in that study were undergraduate college students from Loyola University of Chicago. These
students who viewed the same two films as employed in the present investigation (film of a child who is actually "... emotionally disturbed" and film of a child who is actually normal) were instructed to complete three different questionnaires after each film: the Behavior Description Test, a semantic differential, and the Peterson Problem Checklist. The 1975 study yielded significant results and supported the above hypothesis.

Realizing that a traditional diagnostic label had a powerful biasing effect on an untrained and clinically inexperienced observer who was asked to rate, judge, or interpret a child's behavior, this investigator thought the next step would be to investigate whether a diagnostic label similarly biased a trained and clinically experienced observer. In other research, findings suggest that diagnostic labels may influence clinical judgment and perception of behavior (Gustin, 1964; Langer & Abelson, 1974; Rosenhan, 1973; Sushinsky & Wener, 1975; Temerlin, 1968). It seems obvious from these studies that characteristics such as a clinician's values, personality, mental health training and experience, contact with children, and therapeutic orientation affected the manner in which an event or behavior was perceived, judged, and interpreted. Some of the investigations into such clinician variables were discussed under a previous subheading. It is an interaction of results from this area of research with results
from "social perspective" and "person perception" research on "labeling effects" that has eventuated the current investigation.

**Hypotheses**

First Hypothesis: Filmed children, whether having the actual diagnosis of "emotionally disturbed" or not, when labeled as "disturbed" are perceived (when ratings are summed across the three experimental observer groups; "behavioral" clinicians, "traditional" clinicians, and undergraduates) significantly more negatively than if they are labeled as normal.

Second Hypothesis: The untrained observers (undergraduates) rate the behaviors of the filmed children significantly more negatively than the trained clinician observers ("behavioral" and "traditional" clinicians) rate the same behavior.

Third Hypothesis: "Traditional" clinician observers are influenced to a significantly greater degree by the diagnostic labels imposed on the filmed children than the "behavioral" clinician observers.

Fourth Hypothesis: Untrained undergraduate observers are influenced to a significantly greater degree by the diagnostic labels imposed on the filmed children than the "behavioral" clinicians.
Fifth Hypothesis: Undergraduate observers do not differ significantly from traditional clinician observers in the degree to which they are influenced by the diagnostic labels imposed on the filmed children.

In addition to these hypotheses tested, the researcher had additional concerns which were explored using the data collected. Secondary issues to be discussed are first, whether measures utilizing behavior descriptions closely related to the film stimuli and the more abstract measures less relevant to the specific filmed subjects were consistently sensitive to an imposed diagnosis, and, second, whether the sample of professionals in the "behavioral" and "traditional" categories viewed psychiatric labeling as an aid or a hindrance in their own work.
CHAPTER III

METHOD

Subjects

Subjects were 22 male and 18 female professional therapists or therapists-in-training (clinical psychologists, psychology interns, M.A. psychologists, teacher/psychologists, psychiatric social workers, medical social workers, and social work trainees) from various institutions in the Chicago area.

The sample was obtained by contacting 30 department heads at numerous hospitals, institutes, and universities in the vicinity (Appendix A) which have the reputation of offering either "traditional" or "behavioral" psychological services and training. These contacts were requested to send the investigator a list of therapists who might be willing to fill out the Clinician Questionnaire (Appendix B) utilized in this study to assess therapeutic orientation. Response rate was 30%.

Employing lists of therapists from various facilities as well as lists obtained from more informal sources, 130 explanation letters (Appendix C) and Clinician Questionnaires were distributed to clinicians throughout the area. Response rate was 46%.
Sorting subjects into categories. An analysis of the Clinician Questionnaire by three psychologist-raters followed. An intraclass correlation was performed to determine the interrater reliability of the raters who were asked to sort potential subjects into "traditional," "approaching traditional," "behavioral," and "approaching behavioral" categories (this procedure will be discussed in detail in the Clinician Questionnaire subsection). The intercorrelation of the three raters was found to be significant, \( r_{cc} = .98 \). The reliability of the mean of the three raters was found to be significant, \( r_{cc} = .99 \) (Guilford, 1936).

Those subjects characterized as "definitely behavioral" or "definitely traditional" by at least two raters were considered accurately labeled for the purpose of this investigation. The twenty-five clinicians classified as "behavioral" in this manner and the twenty-six clinicians classified as "traditional" were used as subjects in the present study. Forty-two of these therapists cooperated fully and accurately in the experiment, and two subjects were later randomly eliminated.

The mean age of the professional observers was 33.5, the mean number of years in mental health settings was 8, and the mean number of years of clinical experience with children was 3.7. A two-tailed \( t \) test analysis was performed to compare the average age of the "traditional"
group members with the average age of the "behavioral" group members. The difference in age between groups was not significant, $t (38) = .40, p > .05$. Student's $t$ tests were also performed to establish the comparability of the two groups in terms of total amount of clinical experience and amount of clinical experience with children. It was determined that there was no significant total experience difference between groups, $t (38) = 1.37, p > .05$, and no significant child clinical experience difference between groups, $t (38) = .98, p > .05$. Therefore the investigator concluded that the "behavioral" and "traditional" groups were similar in terms of age and professional experience.

The demographic data presented in Table 1 reveals that the clinician groups were similar in their sex and professional identity makeup, although most "traditionally" oriented clinicians worked in hospital settings, whereas most "behaviorally" oriented clinicians worked in small clinics or laboratories. The "traditional" clinicians were employed at Michael Reese Psychosomatic and Psychiatric Institute, Psychoanalytic Institute of Chicago, Illinois State Psychiatric Institute, and University of Illinois Hospital--Adult and Child Psychiatry Clinics. The "behavioral" clinicians were employed at Institute for Juvenile Research, Illinois Institute for Developmental Disabilities, Department of Psychology at the University of Illinois, University of Illinois Hospital--Child
<table>
<thead>
<tr>
<th></th>
<th>&quot;Behavioral&quot; clinicians</th>
<th>&quot;Traditional&quot; clinicians</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (mean)</td>
<td>31</td>
<td>36</td>
</tr>
<tr>
<td>Years of experience (mean)</td>
<td>6.5</td>
<td>9.5</td>
</tr>
<tr>
<td>Clinical experience with children (mean)</td>
<td>4.6</td>
<td>2.8</td>
</tr>
<tr>
<td>Number of males</td>
<td>14</td>
<td>8</td>
</tr>
<tr>
<td>Number of females</td>
<td>6</td>
<td>12</td>
</tr>
<tr>
<td>Clinicians in hospital settings</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>Clinicians in academic settings</td>
<td>3</td>
<td>--</td>
</tr>
<tr>
<td>Clinicians in special education settings</td>
<td>4</td>
<td>--</td>
</tr>
<tr>
<td>Clinicians in other settings (clinics, labs, etc.)</td>
<td>10</td>
<td>2</td>
</tr>
<tr>
<td>Professional Identity:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Clinical psychologists</td>
<td>8</td>
<td>7</td>
</tr>
<tr>
<td>Psychiatric Social Workers</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Medical Social Workers</td>
<td>--</td>
<td>1</td>
</tr>
<tr>
<td>Psychology interns</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>M.A. psychologists</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Social work trainees</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>Teacher/psychologists</td>
<td>1</td>
<td>--</td>
</tr>
</tbody>
</table>
Undergraduate sample. A comparison experimental group of 20 Introductory Psychology and Developmental Psychology students (10 male and 10 female) from Loyola University of Chicago was also used in the present investigation. The majority of the students were freshmen and sophomores who volunteered to participate. Their mean age was 21. These subjects were a subsample of the 80 students employed in Saper's pilot research on labeling bias (see Pilot Study).

Materials

The films. Two eight-minute, color, 8mm films were used. Film A focused on a normal (has never been involved in psychotherapy and is functioning adequately at home and in school) six year old girl. Film B focused on a five and one-half year old boy who was excluded from the Chicago Public Schools and was attending a special day school for severely emotionally disturbed children affiliated with Loyola University. The combined diagnosis ascribed to him by the Chicago Board of Education and his psychiatrist was: "(1) severe emotional disturbance;
(2) childhood schizophrenia involving pre-psychotic symbiotic ties; (3) mild mental retardation; (4) epilepsy."
For more detailed information on the rationale for utilizing both a normal and disturbed filmed subject, see page 46.

The setting for both films is the Loyola Day School and the grounds of Loyola University. Appropriate releases were obtained. Both children were filmed in similar structured and unstructured activities. They were each filmed alone, with peers (in structured and unstructured activities), and with a teacher or teachers (engaged in both structured and unstructured tasks). Each film was equally divided among these segments. The children were asked to be spontaneous and much of the time they were unaware of the camera or cameraman. However, during indoor filming, especially when they were filmed alone or with a teacher in the room they were cognizant of the filming procedure. Staff members of the Loyola Day School who were familiar with both children informally rated the films as to the accuracy of the footage selected in being representative of their real life behavior. The raters agreed that the behavior of both children in the films was similar to their behavior outside of the experimental situation. To assess the pull or characteristics of the film stimulus itself, the two films were shown to several viewers, professional and inexperienced, who were asked to write descriptions of
the children. See Appendix D for three of these descriptions and see Chapter V for further discussion of the nature of the stimulus. In preliminary research (Saper, Note 6), these films were found to be reliable tools in the discrimination of the effects of a diagnosis.

Clinician Questionnaire. To obtain the appropriate samples for the present investigation, each potential subject was requested by mail to fill out a Clinician Questionnaire (Appendix B). Data from this form revealed the following information about subjects: age, sex, professional background, years of clinical experience, years of clinical experience with children, theoretical orientation of their facility or organization, the orientation of their training programs, and the label with which they would categorize their own brand of therapy or treatment.

To aid in determining whether the potential subjects fell into the "traditional" category or the "behavioral" category, a series of twelve questions about theory of therapy were included. These questions were collected from research by Langer and Abelson (1974), Vardy (1971), and Berzins, Herron, and Seidman (1971), as well as from the writings of Price (1972) and Coleman (1972). Items were constructed employing a four choice Likert scale. Five items were concerned with the degree of agreement or disagreement with an illness perspective of abnormal
behavior, six items with the degree of agreement or disagreement with a psychoanalytic perspective of abnormal behavior, four items with the degree of agreement or disagreement with a learning perspective of abnormal behavior, and three items were concerned with the degree of agreement or disagreement with the social perspective of abnormal behavior (Price, 1972).

Further aids on the questionnaire in determining the category of potential subjects included two questions which ask first for the rank ordering of the psychological factors presented which are most important when practicing therapy and second for the rank ordering of the therapeutic techniques presented which are most important when practicing therapy. The items which the potential subjects were asked to rank include some psychological factors and techniques which are generally thought of as important by most clinicians practicing traditional and behavioral therapies.

Thus, whether a potential subject was utilized or not in the present investigation, depended on how well he or she fit into a "traditional" or "behavioral" therapist category. This was determined subjectively by how they labeled themselves, the theoretical reputation of the facility or organization they were a part of, the orientation of their training program, their answers to the twelve Likert scale items, and their rank ordering of
psychological factors and therapeutic techniques which they considered important or utilize.

In developing this questionnaire, the investigator also hoped to demonstrate the effectiveness of this particular instrument in discriminating these two broad schools of therapy. Other than the A-B Scale (Whitehorn & Betz, 1954) and the TOQ Scale [Fey, 1958; McNair & Lorr (AID Scale) 1964; Sundland & Barker, 1962], very few measures for determining a therapist's orientation can be found in the recent literature. Perhaps this Clinician Questionnaire can be refined and modified so as to be useful in future research.

**Behavior Description Test.** The first test of three tests administered to all subjects was the Behavior Description Test. This measure was developed for exclusive use in the present study. It consists of a series of "positive" and "negative" statements describing the filmed subjects arranged in a Likert-scale type test.

This measuring device was based on a technique used by Langer and Abelson (1974) and Temerlin (1968). They had their subjects write open-ended clinical descriptions about each of their taped subjects including gestures, attitudes, perceived emotional state, interpersonal skills, etc. The measure employed in this investigation was easier to quantify than the open-ended measure. To devise the test, both
films were shown to ten raters who were blind to the experimental hypothesis. Six of the raters were asked to write a character description of each of the children in the style of a "literary narrative." Some of these raters were in the mental health field; others were not. The other four raters were asked to write a clinical report on each child which included theoretical psychological constructs and inferences. These four raters were experienced clinical or developmental psychologists. The reports by the ten raters were surveyed by the investigator who then took statements which either appeared consistently across raters or seemed most representative of the children's film behavior and arranged them in a Likert-scale fashion. Twelve "positive" statements and twelve "negative" statements were included for both children (Appendix E).

The directive to the subjects included in the written instructions on the test was to first go through the statements, marking those which applied to the particular child in the film; and next to rate those statements which applied on a scale from one to three, one being "slightly accurate in describing the child" and three being "quite accurate in describing the child." All mention of the sex of the child was removed from each item so conceivably every item could apply to either child. Two measures were obtained and analyzed from this test. Measure I was the number and strength of positive items chosen by the subject.
Measure II was the number and strength of negative items chosen by the subject. The more "disturbed" the child was perceived to be, the greater the "negative" score (Measure II) and the smaller the "positive" score (Measure I).

**Semantic differential.** The second test administered to all subjects after they had viewed each film was part of a semantic differential devised by Foley in 1971 (Appendix F). The current investigator added one item to the other items ("emotionally healthy-emotionally disturbed") to test the face validity of observer's perceptions of the children's degree of normality or disturbance. Each item in the measure is a bipolar trait and the terms were alternated on a random basis. Some items go from the negative (undesirable) aspect of the trait to the positive (desirable) aspect; others go from the positive to the negative. The traits are rated on a scale from one to six with one being very negative and six being very positive.

The semantic differential (Osgood, 1967) and specifically that scale designed by Foley (1970) is based on a scale used in research by Becker (1960) plus a few additional items. Becker's scale sampled the personality domain outlined by Cattell (1957). Foley used the semantic differential to compare the pre-therapy ratings of a child (both actual and ideal child) by his/her parents and teachers with the post-therapy ratings (cf., Foley, 1970).
Foley's findings, using 48 comparisons for factors, were encouraging for the use of the semantic differential as an adequate measure of behavioral change. She also found through rigorous validation, that the total score on the semantic differential, which is the sum of all the item ratings, is an informative statistic. When the total ratings on the differential for the "disturbed" children (those experimental groups of children in therapy with experienced, untrained, and briefly trained therapists) were compared to the total ratings for the "normal" children (a control group of 50 children matched by age to the experimental groups) the mean total scores and standard deviations were as follows: Clinic Population Mean Total Score-253.43, SD-33.21; Normal Population Mean Total Score-304.13, SD-32.01. (The semantic differentials were completed by each child's mother, father, and teacher. The total scores were then obtained and averaged--the higher the score, the more desirable was the child's behavior.) There was a significant difference, $t(98) = 6.90-10.05$, $p < .001$, on all measures of the "actual" child between the ratings of normal children and the disturbed clinic population. Thus, in Foley's study, the semantic differential discriminated between "normal" and "disturbed" children. Overall, "disturbed" children were rated more "negatively" than normal children. This fact is most important for the present investigation since the total
score rather than factor scores, is the statistic which was considered.

Foley's semantic differential with the addition of the one item mentioned previously, was administered to a pilot population of sixty undergraduate Mundelein College (Chicago, Illinois) students in a "Theories of Personality" class. They were requested to answer each item as it pertained to "the average child" in their opinion. The mean for each item was obtained. Those fifty items with means at either extreme on the one to six scale were the items used in the semantic differential presented to the subjects. The criterion key, based on the numerical value of those fifty items, was used to arrive at a total score of "adjustment." The other items were statistically judged to be ambiguous and of little use for the present experimental groups. Foley (Note 3) suggested that if this current research utilized only a total "adjustment" score and not separate factor scores, then ambiguous items could be safely and statistically eliminated from that total score.

Peterson Problem Checklist. The third test administered to the subjects in this investigation was the Peterson Problem Checklist (1958). This questionnaire (Appendix G) was devised from 20 Cattell-type bipolar scales (Cattell, 1957) which have fairly precise behavioral descriptions. Peterson scored these scales for Cattell's
two largest factors, adjustment and extraversion. The adjustment factor consists of traits such as patient, persevering, mannerly, good natured, calm, responsible, not jealous, cooperative, scrupulous, trusting. The extraversion factor consists of traits such as frank, happy-go-lucky, energetic, friendly, bold, cheerful, assertive, gregarious, composed, prefers companions of the opposite sex. Peterson's Problem Checklist grew from items in these two factors.

The subject in this study circled 0 (no problem), 1 (a mild problem), or 2 (severe problem) if he or she perceived or "guessed" that the statement could apply to the filmed child. The written instructions stated that subjects should "use their imaginations to predict or extrapolate answers from the child's filmed behavior." The total score is the degree of disturbance or maladjustment. The lower the child's total score, the more favorable is the rater's perception and expectation of the child's current and future behavior.

Procedure

Two psychologists and the investigator rated the 60 therapists' Clinician Questionnaires which had been returned and placed them into four categories on two ordinal scales. On the "traditional" scale category one was therapists with a "definitely traditional" orientation.
Category two was therapists who "approach a traditional" orientation. On the "behavioral" scale category one was therapists with a "definitely behavioral" orientation. Category two was therapists who "approach a behavioral" orientation. The following definitions for "traditional" and "behavioral" orientations were provided for the raters reviewing the Clinician Questionnaires. These characterizations were designed by the investigator primarily for this research with the purpose of discriminating between two large groups of clinicians; they should not be considered as having construct validity:

Traditional Orientation: Someone using the psychoanalytic or illness perspective in treating mental disturbances. These perspectives of abnormal behavior are the oldest and most traditional. Some well known writers with these perspectives are, Freud, Rappaport, Ausubel, Meehl, and Kraepelin. Both perspectives come from the medical profession. Deviant behavior is termed pathological and is classified on the basis of symptoms or etiology (the official medical nosology of the APA is generally employed). Diagnosis is of great importance for determining therapy, identifying syndromes, and getting information on prognosis. The psychoanalytic perspective is heavily involved with unconscious psychological processes, intrapsychic conflicts of motives or drives, and the developmental nature of man. The illness perspective is that abnormal behavior is the product of an illness, a compensatory reaction to an organic defect, or a combination of these. Many therapists with a "traditional" orientation employ a "medical model" approach to psychotherapy. They believe that maladaptive behavior cannot be treated directly because it results from underlying causes.
Behavioral Orientation: Someone using the learning or social perspective in treating mental disturbances. These perspectives are fairly recent metaphors developed to understand and treat deviant behavior. Some well known writers with these perspectives are, Skinner, Bandura, Szasz, Scheff, and Becker. The learning perspective views abnormal behavior as learned behavior. The observable behavior is the disorder rather than some underlying state of affairs. The patterns of abnormal behavior are generally explained by identifying sources of reinforcement in the individual's environment. Global personality labels or medical diagnoses are seldom utilized. The social perspective views mental illness as something ascribed to people as a function of the definition given certain types of acts by certain audiences. The social context and the ascription process for behavior is crucial. There is little interest in the etiology of the deviant behavior in question. It is important to consider that deviant role taking has generally been reinforced by society and the labeled deviant is often punished for attempts to return to conventional roles. Social labeling is the most important factor in establishing an individual in a career of chronic deviance.

Those forty-two therapists who were placed in the "definitely traditional" or the "definitely behavioral" categories by at least two of the three raters on the basis of their Clinician Questionnaires were recontacted by mail and by telephone and asked to participate in the second phase of the present investigation. Times were established when the subjects could view the experimental films. The investigator showed the film stimuli to the subjects in sessions consisting of one to eleven participants. The films were always shown to the subjects at their own facility or one near by, at their own convenience.
Clinician subjects were divided into eight counter-balanced groups prior to the film showings. There were five therapists in each group. Four groups were composed of "traditional" therapists and four groups were composed of "behavioral" therapists. Since the diagnostic labels of the filmed children were not given verbally, it was of no consequence that sessions often included members of two or more predetermined experimental groups.

Counterbalancing for film order effects was unnecessary as was demonstrated in the investigator's pilot research (Saper, Note 6). For the main effect of Imposed Diagnosis there had been insignificant order effects, overall.

The investigator began each session with an explanation that he was interested in how clinicians observe and rate behavior of children. The first eight-minute film was shown at sessions in the order described below with the following written instructions on the first page of the test packets (Appendix H). After the film was shown, each subject completed the test packet consisting of the measures described previously. Written instructions for each test were included in the packet. The same procedure was followed for the second film.

To groups I ("traditional" therapists) and V ("behavioral" therapists), the film of the actually
normal child was shown first and they were given the following instruction:

Instruction 1: The child in the short film you are about to see is a normal six year old girl who was filmed while visiting a special day school at which her father is an administrator. The day school for emotionally disturbed children is affiliated with a local public school in Chicago. 

These groups saw the film of the actually disturbed child second and were given the following written instruction:

Instruction 2: The child in the short film you are about to see is a normal 5½ year old boy who attends a parochial school in Chicago. One of his sisters is a paraprofessional at a special school in the area. He was filmed on one of his vacation days while visiting her as a guest at this school.

To groups II ("traditional" therapists) and VI ("behavioral" therapists), the film of the actually normal child was shown first and they were given the following written instruction:

Instruction 3: The child in the movie you will be seeing next is a six year old girl who has been excluded from the Chicago public schools and attends a special school for severely emotionally disturbed children in the area. She is being filmed at this school. She has been given the combined diagnosis by a psychologist working for the Board of Education and her therapist of: mental retardation and severe emotional disturbance produced by a symbiotic psychosis.

These groups saw the film of the actually disturbed child second and were given the following written instruction:
Instruction 4: The child in the film you will be viewing next is a 5½ year old boy who was recently tested by the Bureau of Child Study in Chicago where he was given the following diagnosis by the psychologist and psychiatrist who saw him; severe childhood schizophrenia involving pre-psychotic symbiotic ties; mild mental retardation; and epilepsy. He is being filmed at the special school for emotionally disturbed children which he attends.

To groups III ("traditional" therapists) and VII ("behavioral" therapists), the film of the actually normal child was shown first and they were given written instruction, number 1. These groups saw the film of the actually disturbed child second and were given written instruction, number 4.

To groups IV ("traditional" therapists) and VIII ("behavioral" therapists), the film of the actually normal child was shown first and they were given written instruction, number 3. These groups saw the film of the actually disturbed child second and were given written instruction, number 2.

All subjects were repeatedly assured of their anonymity. The entire experimental session lasted one hour with each small group of subjects. Debriefing was accomplished via a short discussion of the purposes of this investigation after the experimental procedure. At this time subjects' comments on issues they felt the investigation was concerned with and which filmed child they felt was
actually disturbed were elicited. All participants received a copy of the results through the mail at the conclusion of the data analysis.
CHAPTER IV

RESULTS

This investigation sought to determine whether the imposition of a psychiatric diagnostic label on a child biases the perception and rating of that child's filmed behavior. Of particular interest was the way in which an observer's theoretical orientation or training might interact with such a labeling bias phenomenon. To review, the following hypotheses were offered for evaluation:

First Hypothesis: Filmed children, whether having the actual diagnosis of "emotionally disturbed" or not, when labeled as "disturbed" are perceived (when ratings are summed across the three experimental observer groups; "behavioral" clinicians, "traditional" clinicians, and undergraduates) significantly more negatively than if they are labeled as normal.

Second Hypothesis: The untrained observers (undergraduates) rate the behaviors of the filmed children significantly more negatively than the trained clinician observers ("behavioral" and "traditional" clinicians) rate the same behavior.
Third Hypothesis: "Traditional" clinician observers are influenced to a significantly greater degree by the diagnostic labels imposed on the filmed children than the "behavioral" clinician observers.

Fourth Hypothesis: Untrained undergraduate observers are influenced to a significantly greater degree by the diagnostic labels imposed on the filmed children than the "behavioral" clinicians.

Fifth Hypothesis: Undergraduate observers do not differ significantly from traditional clinician observers in the degree to which they are influenced by the diagnostic labels imposed on the filmed children.

The Effect of Diagnostic Labels on Observers' Interpretation Of Behavior

Every subject's total score on each of the four measures discussed in the Methods chapter was obtained and converted to a standard $z$-score. The $z$-score transformations made the scores on the four measures comparable. It should be noted that the $z$-score transformation rules out any main effects for type of measure ($M$) as the mean of each $z$-score distribution is zero. The main effect of orientation ($O$) and the $O \times M$ interaction is interpretable, however. The means and standard deviations for the transformed scores of the various groups are reported in Table 2.
Table 2
Means and Standard Deviations of the Transformed Ratings
On All Four Measures for All Experimental Groups

<table>
<thead>
<tr>
<th>MEASURE</th>
<th>Actual Diagnosis--Abnormal</th>
<th>Actual Diagnosis--Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>Imposed Diagnosis--Abnormal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td>.26</td>
<td>.15</td>
</tr>
<tr>
<td>SD</td>
<td>.98</td>
<td>1.06</td>
</tr>
<tr>
<td>Imposed Diagnosis--Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>x</td>
<td>.45</td>
<td>-.48</td>
</tr>
<tr>
<td>SD</td>
<td>.78</td>
<td>.79</td>
</tr>
</tbody>
</table>
Table 2 - Continued

<table>
<thead>
<tr>
<th>Measure</th>
<th>Actual Diagnosis--Abnormal</th>
<th>Actual Diagnosis--Normal</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>I</td>
<td>II</td>
</tr>
<tr>
<td>Imposed Diagnosis--Abnormal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{x}$</td>
<td>.88</td>
<td>.95</td>
</tr>
<tr>
<td>SD</td>
<td>.40</td>
<td>1.20</td>
</tr>
<tr>
<td>Imposed Diagnosis--Normal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\bar{x}$</td>
<td>-.30</td>
<td>-.70</td>
</tr>
<tr>
<td>SD</td>
<td>1.06</td>
<td>.31</td>
</tr>
</tbody>
</table>

Traditional Orientation Group

| Imposed Diagnosis--Abnormal |   |    |     |     |    |    |     |     |
| $\bar{x}$             | .80| 1.47| .68 | 1.35| -.16| .82 | .50  | 1.08 |
| SD                    | .48| 1.03| .98 | .82 | .76 | .83 | .98  | .65  |

Undergraduate Group

| Imposed Diagnosis--Normal |   |    |     |     |    |    |     |     |
| $\bar{x}$             | .35| .19| -.06| .56 | -.98| -.43| -.83  | -.50  |
| SD                    | .76| .83| 1.03| 1.01| .76 | .32 | .49  | .85  |
The resulting transformed data were then analyzed using a $3 \times 2 \times 2 \times 4$ factorial design with subjects nested within orientation. The factors were orientation ("behavioral" clinician observers, "traditional" clinician observers, and undergraduate observers), actual diagnosis ("severely emotionally disturbed . . ." and normal), imposed diagnosis ("severely emotionally disturbed . . ." and normal), and measure (number and strength of positive behavior descriptions, number and strength of negative behavior descriptions, semantic differential, and Peterson Problem Checklist). The results of the analysis of variance for the combined transformed ratings on all four measures are reported in Table 3.

The results support Hypothesis 1 which stated that filmed children, whether actually disturbed or not, when labeled as "disturbed" were perceived (when ratings are summed across the three experimental observer groups) significantly more negatively than if they were labeled as normal. Inspection of Table 3 indicates that the $F$ value for the main effect of Imposed Diagnosis (I) was significant when the transformed ratings on the four measures were combined and summed across all observer groups, $F (1,144) = 103.47, p < .001$. Examination of the cell means indicates that the children when labeled "disturbed" were rated more negatively than when labeled "normal."
Table 3
Analysis of Variance Summary Table for Combined Ratings on All Four Measures

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>dF</th>
<th>Mean Square</th>
<th>F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orientation (O)</td>
<td>2</td>
<td>11.65</td>
<td>17.81***</td>
</tr>
<tr>
<td>Actual Diagnosis (A)</td>
<td>1</td>
<td>31.18</td>
<td>44.78***</td>
</tr>
<tr>
<td>Imposed Diagnosis (I)</td>
<td>1</td>
<td>72.06</td>
<td>103.47***</td>
</tr>
<tr>
<td>Measure (M)</td>
<td>3</td>
<td>.26E-01</td>
<td>.04</td>
</tr>
<tr>
<td>0 x A</td>
<td>2</td>
<td>1.84</td>
<td>2.81</td>
</tr>
<tr>
<td>0 x I</td>
<td>2</td>
<td>10.91</td>
<td>16.67***</td>
</tr>
<tr>
<td>A x I</td>
<td>1</td>
<td>.36</td>
<td>.52</td>
</tr>
<tr>
<td>0 x M</td>
<td>6</td>
<td>3.68</td>
<td>5.63***</td>
</tr>
<tr>
<td>A x M</td>
<td>3</td>
<td>1.41</td>
<td>2.03</td>
</tr>
<tr>
<td>I x M</td>
<td>3</td>
<td>.63</td>
<td>.90</td>
</tr>
<tr>
<td>0 x A x I</td>
<td>2</td>
<td>3.27</td>
<td>5.00**</td>
</tr>
<tr>
<td>0 x A x M</td>
<td>6</td>
<td>.20</td>
<td>.30</td>
</tr>
<tr>
<td>0 x I x M</td>
<td>6</td>
<td>.33</td>
<td>.51</td>
</tr>
<tr>
<td>A x I x M</td>
<td>3</td>
<td>1.13</td>
<td>1.62</td>
</tr>
</tbody>
</table>

Error Term 1 (For A, I, M, A x I, A x M, I x M, A x I x M) 144  .70

Error Term 2 (For 0, O x A, O x I, O x M, O x A x I, O x A x M, O x I x M, O x A x I x M) 288 .65

** p < .01

*** p < .001
It should also be noted that when total ratings across groups of the filmed boy who was actually disturbed are compared with the total ratings of the filmed girl who was actually normal, there is a significant main effect for Actual Diagnosis (A), $F(1, 144) = 44.78$, $p < .001$. Examination of the cell means demonstrates that behavior of the filmed child who was actually disturbed was perceived and rated more negatively than the behavior of the filmed child who was actually normal regardless of the diagnostic label imposed on them.

The results of probing the significant Orientation (O) main effect, $F(2, 288) = 17.81$, $p < .001$, with the Duncan's new multiple-range test (Edwards, 1965) support Hypothesis 2 which stated that the undergraduate observers (untrained) would rate the behavior of the filmed children significantly more negatively than the trained clinicians rate the same behavior. Specifically, the undergraduates perceived the filmed children, regardless of label, significantly more negatively than the behavioral clinicians and more negatively than the traditional clinicians ($p < .05$). There was no significant difference between the two trained clinician groups in the total amount of pathology they perceived in the two films. The shortest significant ranges are as follows: $\alpha = .05$; $R_2 = .35$; $R_3 = .37$.

The results of probing the significant $O \times I$ interaction, $F(2, 288) = 16.67$, $p < .001$, with the Duncan's test
support Hypothesis 4 and partially supports Hypotheses 3 and 5; which stated that traditional clinician observers would be significantly more influenced than behavioral clinician observers by the diagnostic labels imposed on the filmed children (3); undergraduate observers would be significantly more influenced than behavioral clinician observers by the diagnostic labels imposed on the filmed children (4); and undergraduate observers would not differ significantly from traditional clinician observers in the degree to which they are influenced by the imposed diagnostic labels (5). The shortest significant ranges from the Duncan's test are as follows: $\alpha = 0.05; R_2 = 0.50, R_3 = 0.52, R_4 = 0.54, R_5 = 0.55, R_6 = 0.56$.

The additional results of probing the significant $0 \times A \times I$ interaction, $F(2, 288) = 5.00, p < 0.01$, lead to a clarification of the relationships among an observer's orientation, the actual pathology of the filmed children, and the diagnostic labels imposed on these children. The "traditional" clinician observers and the undergraduate observers rated the behavior of the disturbed child who was labeled as such significantly more negatively than the "behavioral" clinician observers rated it ($p < 0.05$). They also rated that behavior significantly more negatively than when the disturbed child was labeled normal, lending support to Hypotheses 3, 4, and 5 ($p < 0.05$). However, when the normal filmed child was labeled "disturbed,"
the undergraduate observers were significantly biased by the label ($p < .05$) but the "behavioral" group was not, lending support to Hypothesis 4, and the "traditional" group was not, which does not support Hypotheses 3 or 5. In other terms, the undergraduate observers were influenced by imposed diagnostic labels whether pathology was actually present or not, the "traditional" clinician observers were influenced by imposed labels only when pathology was actually present, and the "behavioral" clinician observers were not influenced by imposed diagnostic labels whether pathology was actually present or not. The shortest significant ranges from the Duncan's test are as follows: $\alpha = .05$; $R_2 = .70$, $R_3 = .74$, $R_4 = .77$, $R_5 = .79$, $R_6 = .86$, $R_7 = .81$, $R_8 = .82$, $R_9 = .83$, $R_{10} = .84$, $R_{11} = .84$, $R_{12} = .85$.

A comparison of the two film stimuli in the probing of the same $O \times A \times I$ interaction indicates that none of the three observer groups differentiated to a significant degree in their perceptions and ratings, between the film of the actually disturbed boy and the film of the actually normal girl when the children were both given the same diagnosis. The "behavioral" clinician observers did not differentiate between the film stimuli regardless of the imposed label.

**Secondary Concerns**

This investigator was additionally interested in how consistently measures utilizing behavior frequencies
and behavior descriptions closely related to the filmed subjects (Measures I and II) and the more abstract measures less relevant to the specific film subjects (semantic differential and Peterson Problem Checklist) were sensitive to an imposed diagnostic label. The summarized results of analyses of variance performed on each of the four measures individually are reported in Table 4. The results support the notion that all four measures, although they are each tapping qualitatively different types of responses about the filmed stimuli, are sensitive to an Imposed Diagnosis main effect as well as the actual pathology or adjustment of the children in the films (Actual Diagnosis main effect). An $0 \times 1$ interaction is demonstrated on Measures I, II, and IV. Probing this significant interaction on each measure indicates that on Measure I (number and strength of positive behavior descriptions), the "traditional" clinicians and the undergraduate group observed in the children carrying the imposed diagnosis of "disturbed" fewer positive behaviors than they observed in the same children carrying the imposed "normal" label, $F(2, 72) = 5.19, p < .01$. The shortest significant ranges from the Duncan's test are as follows: $\alpha = .05$; $R_2 = .53$, $R_3 = .57$, $R_4 = .58$, $R_5 = .61$, $R_6 = .62$. On Measure II (number and strength of negative behavior descriptions) the "traditional" clinicians and the undergraduate group observed in the children carrying the imposed diagnosis of "disturbed" more negative behaviors than they
Table 4

ANALYSES OF VARIANCE F VALUES FOR EACH MEASURE INDIVIDUALLY

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>Measure I (Number and Strength of positive behavior descriptions)</th>
<th>Measure II (Number and strength of negative behavior descriptions)</th>
<th>Measure III (Semantic differential)</th>
<th>Measure IV (Peterson Problem Checklist)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Orientation (O)</td>
<td>.99</td>
<td>13.97**</td>
<td>.63</td>
<td>24.06***</td>
</tr>
<tr>
<td>Actual Diagnosis (A)</td>
<td>30.65***</td>
<td>11.57**</td>
<td>4.86*</td>
<td>5.94**</td>
</tr>
<tr>
<td>Imposed Diagnosis (I)</td>
<td>15.29***</td>
<td>33.71***</td>
<td>24.00***</td>
<td>36.52***</td>
</tr>
<tr>
<td>O \times A</td>
<td>1.84</td>
<td>.38</td>
<td>.10</td>
<td>1.56</td>
</tr>
<tr>
<td>O \times I</td>
<td>5.19**</td>
<td>4.59*</td>
<td>2.41</td>
<td>5.90**</td>
</tr>
<tr>
<td>A \times I</td>
<td>.28</td>
<td>5.43**</td>
<td>.08</td>
<td>.00</td>
</tr>
<tr>
<td>O \times A \times I</td>
<td>.16</td>
<td>1.80</td>
<td>1.37</td>
<td>2.93</td>
</tr>
</tbody>
</table>

* p < .025
** p < .01
*** p < .001
observed in the same children carrying the imposed "normal" label, $F(2, 72) = 4.59, p < .025$. The shortest significant ranges from the Duncan's test are as follows: $\alpha = .01; R_2 = .60, R_3 = .62, R_4 = .64, R_5 = .65, R_6 = .66$. On Measure IV (Peterson Problem Checklist) the "traditional" clinicians and the undergraduate group expected, extrapolated, and predicted more negative behaviors from the children with the imposed diagnosis of "disturbed" than from the same children when labeled "normal," $F(2, 72) = 5.90, p < .01$. The shortest significant ranges from the Duncan's test are as follows: $\alpha = .01; R_2 = .60, R_3 = .62, R_4 = .64, R_5 = .65, R_6 = .66$.

On all three of these measures there was no significant difference in ratings from the "behavioral" group between the filmed children when labeled "normal" and the filmed children when labeled "disturbed." Only the semantic differential, Measure III, picked up a significant labeling bias effect ($F(1, 72) = 24.0, p < .001$), across all three groups with no significant differentiation between them.

The probing of the significant $0 \times I$ interactions on the separate measures also indicates that the undergraduate observers, in general, perceived more negative behaviors in the films than the trained clinicians if the child in the film had the imposed label, "disturbed" ($p < .05$). The "traditional" clinicians perceived significantly more
negative behaviors in the same experimental situation than did the "behavioral" clinicians ($p < .05$). Yet at the same time, the "behavioral" clinicians generally perceived significantly fewer positive behaviors in the filmed children if they were called "normal" than did the "traditional" clinicians ($p < .05$). In addition, the undergraduates had significantly more negative expectations, extrapolations, and predictions about the behavior of the children regardless of the imposed label, than had both groups of clinicians ($p < .05$). When the filmed children were labeled "disturbed" the "traditional" group had significantly more negative expectations, extrapolations, and predictions of the children's behavior than had the "behavioral" group ($p < .05$).

The other secondary concern dealt with in this investigation was whether professionals in the "behavioral" and "traditional" categories viewed psychiatric labeling as an aid or a hindrance for themselves and their patients. The two experimental groups did differ significantly with a one-tailed $t$-test analysis in their answer to the labeling item on the Clinician Questionnaire. That is, when asked to rate the helpfulness to therapists of official APA psychiatric diagnoses on a scale from 1 (a hindrance) to 4 (very helpful), "behavioral" clinicians had a mean rating of 1.7 and "traditional" clinicians had a mean rating of 2.1, $t(38) = 1.84$, $p < .05$. "Traditional" clinicians found the utilization of traditional diagnostic
labels less of a hindrance than "behavioral" clinicians found them, although both groups viewed such labeling on the hindrance, rather than the helpful, end of the continuum.
CHAPTER V

DISCUSSION

The applied mental health field, to a large degree, is crucially dependent upon the clinical judgment, accuracy, and efficacy of its professional and paraprofessional membership. It is essential that the language used and labels ascribed during the rendering of mental health services do not in any way function as a deterrent to the rights and freedoms of the individual consumer. The current study presents evidence that the imposition of traditional psychiatric diagnostic labels on a child biases the perception of and response to that child's filmed behavior, whether such behavior is pathological or not. Specifically, the data presented for evaluation previously demonstrate and support Hypotheses 1, 2, and 4, and partially support Hypotheses 3 and 5. The discussion will follow the order in which the hypotheses were stated on page 75 and will then turn to considerations of secondary issues, methodological concerns, future research ideas, and implications of this investigation.

Labeling Bias: General

When all observer groups were combined in the statistical analysis, a labeling bias effect was demonstrated.
Both children, whether actually disturbed or not, were generally perceived more negatively or more pathologically when they carried the imposed diagnosis of "emotionally disturbed . . ." than when they were called "normal" (Hypothesis 1). One example of this was an observer's tendency to respond on the semantic differential that the child in the film was slightly to moderately "aloof" when diagnosed disturbed; whereas the observer responded that the same child was slightly to moderately "responsive" when labeled "normal." These same observers viewed certain specific behaviors (whether those behaviors were actually adaptive or maladaptive) as "deviant" when the child was labeled "severely emotionally disturbed" and as "typical" when the child was labeled "normal." For example, the same filmed incident indicated on the Behavior Description Test (Appendix D) was often described as "the child . . . appeared to be hallucinating or at least losing contact with events and circumstances around him or her" when the "emotionally disturbed . . ." diagnosis was imposed, or described, "like many kids this child makes funny faces . . ." when the "normal" label was imposed.

The support for Hypothesis 1 indicates that a diagnostic label imposes a response set on the observer which makes his/her judgments inaccurate. Cronbach (1964) categorized this measurable component of inaccuracy in
social perception as "elevation." One must interpret this result with caution. As will be discussed with the remarks on Hypothesis 3, summing over the observer groups does indicate a general labeling bias phenomenon; however, the three observer groups are differentially subject to such a phenomenon.

Perceived Pathology: A Function of Training and Actual Pathology

The undergraduate observers, all of whom were untrained in the field of mental health, rated both children more negatively than the two groups of trained clinicians (Hypothesis 2). This finding suggests that the undergraduates perceived and judged more pathology present in the behavior of the filmed children than the mental health professionals. Such results do not seem surprising. Perhaps the clinicians with more experience working therapeutically with children and operating in the mental health field in general have been trained or at least have learned to be more cautious and objective in their judgments of pathology. Another important factor might be the difference in age of the observers. The undergraduates, on the average, were twelve years younger than the professional observers. Thus, they are less likely to have children of their own. Since they lack child references for behavior other than themselves or siblings, they saw the behavior of the filmed children as more primitive and strange when compared to their own.
The undergraduates, although relatively untrained in the field of mental health, probably have some academic knowledge of psychology and the meaning of some psychological diagnoses. Students such as these are often found in mental health settings as paid paraprofessionals or volunteers. Results of examining Hypothesis 2 indicate that without the proper training of the nonprofessional staff, child clients in these settings may be judged more pathologically than is warranted. From Rosenthal's (1973) study of school children, it was found that teachers responded to children they thought had less potential than others with less warmth and openness, and gave them fewer opportunities to achieve (see p. 25). A possible implication from the present study is that nonprofessionals working with children carrying APA diagnoses may relate to these children in a biased manner reminiscent of Rosenthal's teachers.

It is interesting to note that the results also indicate that the filmed boy (actually disturbed) was generally rated more negatively than the filmed girl (actually normal) when the ANOVA cells were collapsed over observer groups, imposed diagnosis, and measure. It might appear that, by and large, subjects could discriminate real differences in mental health between the two children. However, in the two situations when both children carried the same diagnosis none of the three
groups rated the children significantly differently from each other. It should be remembered that other elements in the films which differentiate the boy from the girl confound these statements.

In "person perception" terms, the accuracy of observers' perception in this project generally hinged on the demand characteristics of the stimuli and the inherent characteristics of the observer. Some of the salient variables in the stimuli used were as follows: one filmed subject was a boy and the other was a girl; the girl was better dressed than the boy; the girl looked slightly older than the boy; the films are silent and both children appeared to be talking at various times (in actuality, the boy's utterances were not conversational, whereas the girl's were); both subjects were filmed in a special school setting (observers were told this), the girl was a stranger to the other children in the film, while the boy was not; both children were filmed in structured and unstructured situations, alone, with peers, and with adults; and the boy had more contact with controlling adults in the film than the girl since he actually was a special student at the day school. The film clips of both children were judged to be representative samples of the childrens' behavior by adults who had formal contact with them. If the reader is interested in examples of descriptions of the filmed children from experienced professional and
inexperienced undergraduate observers who were given no diagnostic labels, three such descriptions can be found in Appendix D.

Because of the cues available that one responds to in meeting the demand characteristics of the films (e.g., sex, age, clothing, etc.), one might conclude that these variables could be confounding the effect an imposed diagnosis has on the perception of the observers. In the current study the labels affixed to the two filmed children obviously differ, but so do other characteristics of the children such as the variables just mentioned. Therefore any direct comparisons of the "normal" subject and the "disturbed" subject are ambiguous since they may be based on actual differences in behavioral abnormalities, on other cues such as sex, or a combination of both of these. However, this type of comparison per se is not related to the primary concern of this study. The main comparisons which this investigator has been focusing on are not affected by cues other than the imposed diagnostic label since each film subject is compared with him/herself.

In addition to the inherent characteristics of the stimuli, the inherent characteristics of the observers are also crucial variables which will be discussed under the next subheading.
"Traditional" vs. "Behavioral" vs. Undergraduate Observers

Langer and Abelson (1974) suggested that diagnostic labels provide one vehicle for organizing the input surrounding any situation or individual. Diagnoses "serve as categories or sets that in addition to structuring the previous input, determine what further information is attended to" (p. 8).

By assigning different labels to the filmed children, different types of observers may be led to view them and attend to further information about them in disparate ways. In the current investigation the "behavioral" and "traditional" observers were differentially influenced by the diagnostic labels imposed on the children (Hypothesis 3). The "traditional" clinicians were biased by a psychiatric diagnosis when it was imposed on the child who actually manifested some pathology. Thus, when a severely disturbed child was characterized as "normal," a "traditional" therapist tended to perceive and interpret his behavior in a much more positive light than if the same child was accurately characterized. On the other hand, "behavioral" clinicians tended to remain uninfluenced or not biased by a psychiatric diagnosis in their perceptions and interpretations of behavior, when a diagnosis was imposed on either a disturbed or normal child.

This finding does not offer an explanation as to why "behavioral" therapists seem to be more immune to a
labeling effect than "traditional" therapists. The "behavioral" therapists may be focusing so strongly on observable behaviors in the films that they tend to ignore previewing information such as the imposed label. Another less plausible possibility is that they are unfamiliar with the definitions, implications, and utilization of psychiatric/medical labels, so they discount them, or they take the diagnoses into consideration, process them, and reject their relevance to the children's filmed behavior because their training emphasizes the use of behavior descriptions and frequencies rather than global personality diagnoses. It should be mentioned again that the behavior therapists tended not to perceive or interpret the filmed boy's behavior as significantly more pathological than the filmed girl's behavior, regardless of the imposed label. Thus, the "behavioral" clinicians may be either insensitive to the film stimuli, or they may be less interested in selecting pathological behaviors than in selecting behaviors which indicate adjustment or strengths, regardless of the mental status of the child; or other characteristics of the films may be confounding the mental status difference between the two filmed children.

The present results partially support Hypothesis 3, that the "traditional" clinician observers differ significantly from the "behavioral" clinician observers in
the degree to which they are influenced by the diagnostic labels imposed on the filmed children. However, as with the "behavioral" group, this finding does not offer an explanation as to why the "traditional" group demonstrated more of a labeling effect than the "behavioral" group (demonstrated only with the boy film stimulus). The "traditional" group may be employing the imposed diagnostic information to filter the totality of cues in the film, or they may be projecting their general "stereotype," as discussed by Cronbach (see p. 24), of what the mentally ill or mentally healthy child is like on the filmed boy rather than a real evaluation of his behavior. There may be several possibilities for why the "traditional" clinicians succumbed to the labeling effect when viewing the film of the boy and not when viewing the film of the girl. Perhaps the interaction of actual pathology with a pathological diagnosis fits best into their general notion of mental illness, whereas an interaction of adjustment with a pathological diagnosis forces them to reject the pre-film information. A label of "normal" imposed on either of the children appears to force them to reject, in their perceptions or interpretations, incidents of pathology which may be present. However other characteristics which differentiate the two stimuli (as mentioned previously) could also be operating. For example, perhaps "traditional" clinicians are more biased
by labels when they are relating to boys than when relating to girls.

It is noteworthy that the "traditional" clinicians, in general, rated the behavior of the disturbed child when he was labeled as such significantly more negatively than did the "behavioral" clinicians. This may be caused by their general illness or medical perspective of abnormal behavior; or they may have been more sensitive to actual maladaptive behavior which the boy evidenced in the film than were the other clinicians. In either case, it is likely that a traditional, analytic, or medical perspective of mental illness tends to focus more on pathological behavior than a behavioral or social perspective, and when pathology was actually present it was shown that clinicians with a "traditional" orientation were more influenced by diagnoses, yet also better able to pick up real pathology, than were clinicians with a "behavioral" orientation.

It is evident that any type of theoretical orientation may be subject to some types of bias or influence. However, the partial support of Hypothesis 3 might suggest that the "behavioral" group was better able to avoid the type of bias of perceptions of behavior and clinical judgments that a diagnostic label imposed on a child can introduce than the "traditional" group. Another possibility for such a finding might be that the "behavioral" clinicians were counterbiased by diagnostic labels, i.e., they became
insensitive to behavior differences and/or pathology when a diagnosis was presented and tended to mistrust all previewing information.

The undergraduate observers were significantly biased by the imposed labels whether the child was actually disturbed or not. Labeling bias was significantly greater for the untrained undergraduates than for the "behavioral" clinicians (supporting Hypothesis 4), and there was a significant difference in labeling bias between the undergraduates and the "traditional" clinicians when the normal child was labeled "disturbed" (partially supporting Hypothesis 5). It seems possible that mental health training, formal experience with children, and possibly greater maturity affect not only judgments of pathology, as previously discussed, but also how biased one is by diagnoses. Results of this study indicate that a fallacious diagnosis imposed on a child may affect how an inexperienced observer, such as a member of the undergraduate sample, perceives the child's behavior, interprets that behavior, judges the child, responds and relates to the child, expects the child to perform, predicts the child's future performance, and selects behaviors to modify, strengthen, or eliminate.

Secondary Concerns

The investigator had two additional interests:
first, whether measures utilizing behavior descriptions closely related to the film stimuli and the more abstract measures less relevant to the specific filmed subjects were consistently sensitive to an imposed diagnosis; and second, whether the sample of professionals in the "behavioral" and "traditional" categories viewed psychiatric labeling as an aid or a hindrance in their own work. 

The results indicate that all four measures were sensitive to an imposed diagnostic label. The semantic differential picked up a label bias across all groups of subjects with no significant differentiation between them. Perhaps because the semantic differential tapped abstract, global, and possibly ambiguous personality traits of the filmed children, all trained and untrained observers, when using it to judge the children's behavior succumbed to a labeling bias. Another possibility is that the semantic differential tapped attitudes rather than observations and perhaps attitudes toward the filmed children were more readily altered by diagnoses than more objective behavioral descriptions. When measures were more concrete and behaviorally relevant to the film stimuli (Measures I and II), all observers and especially the "behavioral" clinicians were less apt to be influenced by a fallacious label. The "traditional" therapists and the undergraduates were significantly influenced by the imposed labels even when they were rating the children's behavior on such an
actual behavior description instrument. Perhaps they are less accustomed to examining and interpreting specific observable behaviors and responses to stimuli than the "behavioral" therapists.

When the measure dealt with concrete predictions, expectations, and extrapolations for present and future behavior (Measure IV), the "behavioral" clinicians were less apt to be biased by imposed labels than the other observer groups. The "traditional" clinicians and the undergraduates expected, extrapolated, and predicted significantly more negative behaviors for the children with the imposed diagnosis of "disturbed . . ." than for the same children when categorized as normal. Additionally, the "traditional" and "behavioral" clinicians tended to be less severe overall in their outlook for the children's total present and future behavior, regardless of label, than the undergraduates. Once again, this is probably a function of mental health training, formal experience with children, and maturity.

The "traditional" clinicians made more pathological predictions of the children's behavior (Measure IV) and perceived and interpreted more negative behaviors (Measure I) than the "behavioral" clinicians. Once again, this may be a function of an illness perspective of abnormal behavior or an accurate assessment of pathology which may be present in the children's behavior. Since the
"behavioral" clinicians tended to perceive significantly fewer positive behaviors in the filmed children when they were called normal (Measure I), than the "traditional" clinicians, perhaps behavior therapists generally tended to avoid extreme statements about both maladaptive and adaptive behavior. Traditional therapists tended to be more extreme in their ratings. Depending on the bias of the person blindly rating the films or interpreting the results, the behavioral group could be described either as more objective in their behavior ratings or more insensitive to the differences between the films than the "traditional" group.

The findings clearly indicate, however, that all observer groups were significantly affected by labeling bias on an abstract, global trait measure, the semantic differential, whereas on more concrete, behavior-specific measures, at least the "behavioral" group proved to be immune to a labeling bias phenomenon. For this reason, it is heartening that the growth of behavioral approaches in psychotherapy in the last twenty years has stimulated interest in new systems of classification which utilize units smaller than the global personality which is the unit more often assessed for labeling (Scott, 1968). Behaviorists suggest that the reliability of diagnostic labels will be higher if those labels are based on specified, meaningful, discernible behaviors rather than on
total personality. However, Kass and O'Leary (Note 4) found, as was also found in the present study, that even when specific behaviors are observed and rated, observers could be biased by expectations or labels.

To reduce or eradicate this source of error, perhaps individuals can be trained to record behavior in clearly defined categories. In a study by Kent, O'Leary, Digment, and Dietz (Note 5), it was found that raters' expectations of children's disruptive behavior in a classroom affected their "global evaluation" of the extent of behavior change observed, but "specific behavior recording" produced by the same observers after instruction did not show any effect of the expectations. In applied mental health settings, the training of nonprofessional observers and technicians and even professional clinicians in defining target behaviors specifically and thoroughly might seem warranted from the results of the present study, and others mentioned above, to eliminate the biasing effect of diagnoses or imposed expectancies. In conjunction with this, nonprofessional and professional direct service providers in these settings, especially where children are the recipients, should be cautious when responding to or employing traditional, global, psychiatric diagnoses of their clients.

The current study reinforces this necessity for caution in the utilization of diagnostic labels, even if
they are behavioral labels based on objective measures. Although there may be a clear consensus among mental health workers as to whether or not a particular behavior has occurred there is still room for considerable disagreement as to the behaviors employed as criteria for "normality."

The other secondary concern of the investigation was whether "behaviorists" and "traditionalists" differed significantly in how they rated the helpfulness of official APA diagnoses. It was found that "traditional" clinicians approached the utilization of traditional diagnostic labels less negatively than "behavioral" clinicians. It is interesting to note that the clinicians employed in the present study all viewed traditional labeling on the hindrance end of the continuum, although the practice of diagnostic labeling is widespread in hospitals, clinics, and governmental agencies. The finding that "traditional" therapists saw diagnoses as less of a hindrance than "behavioral" therapists is congruent with results of the Langer and Abelson (1974) study. Their behavior therapist group generally scored lower on the item (diagnoses are less of an aid than a hindrance) than their traditional therapist group (diagnoses are more of an aid than a hindrance). One reason the "traditional" group may have succumbed to a label bias in some cases, whereas the behavioral group did not, is that these clinicians are more
comfortable and content with the employment of such labels and generally find them to be of more utility than the "behavioral" clinicians.

Methodological Issues

At this point it is important to understand some of the methodological and ethical criticisms lodged against some of the basic "social perspective" studies previously mentioned (e.g., Gustin, 1969; Langer & Abelson, 1974; Phillips, 1963; Rosenhan, 1973; Sushinsky & Wener, 1975; Temerlin, 1968), and briefly discuss ways in which the current research perhaps improved upon them. Also, an analysis of the limitations of this study as well as problems with the design employed can be profitably discussed at this time. In "diagnostic fallibility" research, "expectancy and person perception" research, and "attitudes toward deviancy" research, the stimuli which observers or subjects are asked to rate, interpret, and respond to are of the utmost importance. As discussed on page 45, investigators in these areas have had subjects rate, respond, and react to written or verbal descriptions of an individual; a tape recording of an individual; a video tape or film of an individual; a photograph or test protocols of an individual; a live individual; etc.

Phillips (1963) employed written descriptions of three hypothetical people--two of whom "suffered" from
forms of "schizophrenia" and one of whom was normal. His 300 subjects reacted to these written stimuli and from these reactions, he gleaned their attitudes about mental illness. A short written character description of a hypothetical person contains so much less behavioral data than an actual live person, that it is hard to say with certainty that Phillips' conclusions about his subjects' attitudes toward mentally ill people are valid in the real world.

Rosenhan (1973) refined this methodology by using "live" stimuli rather than written descriptions, tapes, or films. He had normal stooges fake symptoms and feign insanity at intake interviews, and he collected information on hospital staffs' treatment of pseudopatients labeled "schizophrenic." A limitation of this design, however, is that Rosenhan did not compare the reaction to these stooges with the reaction to people who were actually schizophrenic at intake. One ethical criticism Rosenhan often receives (see Social Perspective section) is that his stooges lied about their complaints and distorted their histories leaving admittance to the hospital as the most plausible, humane avenue open for the intake physicians. In Rosenhan's corollary study he told hospital staffs that pseudopatients would be trying to get themselves admitted to the hospital, and a number of admitting physicians subsequently rated actual patients as
fakers. Unfortunately, no pseudopatients actually went through the intake process so no comparison could be made between reactions to those feigning insanity and those who were insane.

Temerlin (1968) as well as Sushinsky and Wener (1975) had observers react to a veridical taped script interview of a normal person/actor who they were told "looked neurotic, but actually was quite psychotic" by a credible and reputable source (see Social Perspective section). To further expand on the conclusions drawn from his results, it might have been interesting if these researchers had also used a tape of a person who actually was psychotic.

Langer and Abelson (1974) used a video tape of an individual being interviewed as the stimulus to which their subjects responded (see p. 19). This individual was an actor and was in reality, neither participating in a "job interview" nor a "psychiatric interview." Langer and Abelson were unable to compare professionals' reactions to, and diagnosis of, an individual who was actually being interviewed for a job with their reactions to an individual who was actually in therapy, since they used only one videotaped subject as their stimulus.

In the present investigation two filmed subjects were used as stimuli to which observers reacted. One of the filmed subjects as indicated in Chapter III was actually
emotionally disturbed--having been excluded from the public schools and placed in a "special school," and the other subject was actually normal (never involved in psychotherapy and adequately functioning at home and in school). The imposed diagnosis (and only the imposed diagnosis) was manipulated for both filmed subjects. Therefore the responses to the disturbed subject when labeled "normal" could be compared to the responses to the disturbed subject when labeled "disturbed," and similarly for the "normal" subject. Using this type of methodology the investigator will avoid criticisms of "faked histories," "distortion" or "phony symptoms." This methodology, unique to research in the area seems to make the results somewhat more generalizable, and more applicable in the field than results from previous investigations because it avoids many of the criticisms leveled against earlier research.

The general limitation on the stimuli used in the current study is that observers are watching silent films of children rather than live children. Critics could argue that the additional behavioral information which professionals have about, real, live children they are working with is enough to eliminate any biasing effect a diagnostic label might have. One response to such criticism is that professionals who often impose or react to diagnoses of children are sometimes less involved with the actual child than with the child's description by others,
testing protocols, or small samples of the child's behavior that the professional has witnessed. It is possible that silent films such as those utilized in the present study, edited to be representative of the children's everyday behavior, gave the clinician observers more useful information about the filmed subjects than they would normally obtain about their child clients from pretherapy material or from an initial contact.

Other issues did arise because of the nature of the two film stimuli. After each experimental session, the subjects were casually asked which filmed child they thought was actually disturbed. About 60% of the subjects perceived the girl to be less disturbed than the boy. About 50% of the subjects also guessed that the investigation had "something to do with labeling" or the written information given them about each child. There was no significant difference between the reaction of the two clinician groups. Of those subjects, about 75% of them still felt they had been influenced by the labels. Ten percent said they did not feel they had been influenced, and 15% were unsure about whether they had been influenced or not. This finding implies that although clinicians may be aware of the possibility of labeling bias, the majority may still succumb to such biasing.

Since the film of the girl was always shown before the film of the boy in the experimental sessions, critics
may charge that film order interactions will bias the findings. However, in pilot research done by this investigator (Note 6) using the same film stimuli, no significant order effects were demonstrated for the Imposed Diagnosis effect. It is this main effect which was examined in Hypothesis 1. It is possible, although unlikely, that interactional effects such as orientation by imposed diagnosis (examined in Hypothesis 3, Hypothesis 4, and Hypothesis 5) might be confounded by an order effect. The actual main effect in the prior research was significant on three measures when a film was shown first and on one measure when a film was shown second. Thus a possible, although improbable, confounding order effect should be considered when examining any relationships comparing the filmed boy with the filmed girl and interaction between actual diagnosis and other effects. Perhaps, in future research an efficient, practical, and economical method should be devised for totally counterbalancing experimental groups, including counterbalancing for film order, as was done in Saper's 1975 research.

Sushinsky and Wener (1975, p. 82) suggest from their research findings that "labeling bias is extremely general, and probably related to attractiveness of the communicator who gives the suggestion of mental illness ... and other similar variables well documented in social psychology literature." Temerlin (1968) proposed
that labeling bias is an interaction of "professional identity" and prestige suggestion. Thus, the way in which diagnostic information was communicated about the filmed children in the current investigation, as well as the prestige of the communicator, are important methodological issues to confront. Since a general labeling bias effect was demonstrated (Hypothesis 1), it is probable that most group members were attending to the written label suggestions and pre-viewing information. It is also possible that the "behavioral" clinicians did not attend to the label suggestion since they were immune to the labeling bias effect. However, when the members of this group were debriefed after the experimental sessions, all of them stated that they had read and attended to the diagnostic information presented before each film, although many said they rejected the information as being irrelevant. It is probable that the prestige of the communicator of the diagnostic label in the written description of each filmed child was an important factor. The "abnormal" diagnoses imposed on the girl were attributed to a Board of Education psychologist and her therapist. The "abnormal" diagnoses imposed on the boy were attributed to a Bureau of Child Study (Board of Education) psychologist and psychiatrist. The "normal" label imposed on the children was not attributed to any professional, although the girl's father was stated as being an administrator of a special school, and the boy's
sister was stated as being a paraprofessional in a special school. The identity of the communicators of the information, the relatives of the children, and their specific diagnoses were intentionally slightly different since every subject viewed both films and read both of the diagnostic descriptions at the same sitting. However, it was felt that the descriptions were comparable and similar in content and tone. In future research this comparability should be experimentally tested.

Some professional subjects informally discussed with the investigator the notion that the written sources for the children's diagnoses were not always considered very credible communicators of diagnostic information. This was especially true for those clinicians who had a good deal of contact with these institutions and provide service for children referred from them. Interestingly, the undergraduate group demonstrated the most labeling bias and they probably were most impressed with the prestige of the communicators employed in this investigation; therefore they were more apt to believe and be influenced by the diagnostic information than the other subjects. The "behavioral" group demonstrated the least labeling bias and they tended to come into more contact with children, especially in educational and clinic settings than the other clinician group. Thus, they may have had more knowledge of the Board of Education's labeling procedures
than the "traditional" group and perhaps tended to mistrust or ignore diagnostic information coming from that institution. This finding strongly suggests that the labeling bias effect may indeed be an interaction between theoretical orientation, training, and prestige of the labeler, rather than theoretical orientation or training alone.

Methodological problems and possibly assets also include the lack of uniformity in the measures used to tap observers' reactions to the filmed subjects. As demonstrated, some comparisons were made among measures because a $z$-transformation was performed on the data although weights were not assigned. It is in this area where the current investigation differs significantly from most other similar studies. Utilizing three different types of rating devices, the investigator was able to evaluate the degree to which the imposed diagnoses biased observers' expectations for the children's present and future behavior (Peterson Problem Checklist), global impressions of the children's personalities (semantic differential), and specific ratings of actual filmed behavior (Behavior Description Test). No other investigation in the literature employed instruments to tap all three of these areas.
Another methodological issue with which this investigation was concerned was the determination of a therapist's "theoretical framework." Working with the hypothesis that "behaviorally oriented" clinicians would be affected differently than "traditionally oriented" clinicians by imposed psychiatric labels, the investigator had to establish a reliable method to sort professionals into such categories. It has generally been thought that the theoretical orientations clinicians have are products of the interaction between their personalities and their training. Empirically, only a small number of individuals identify totally and forcefully with one ideological system. The majority of mental health professionals are eclectic to varying degrees. However, researchers have always been interested in categorizing the types of psychotherapies practiced, and how they are differentially effective with various types of clients and how practitioners are differentially affected by such factors as: pre-therapy information, post therapy information, client variables, therapist variables, and the interaction of these (e.g., Berzins et al., 1971; Langer & Abelson, 1974; Vardy, 1971; Whitehorn & Betz, 1954).

Several investigators employed self reports in their research such as: "What label would you give to the type of psychotherapy you practice?" or "Which authors shaped your present therapeutic approach?" (Armor &
Klerman, 1968; Berzins et al., 1971). Weiss (1973) was one of several researchers who categorized clinicians according to the theoretical reputation of their training facility. Other researchers had their clinician subjects answer relevant questions to assess their ideology (Langer & Abelson, 1974; Vardy, 1971).

The approach used to categorize the professionals in the current investigation was a combination of the above techniques. However, the development of the investigator's Clinician Questionnaire was not statistically rigorous. A significant interrater reliability on the instrument, $r_{cc} = .98$, suggests that the tool is a reliable one for the purposes of the current study, however, until it has been subject to further refinement, it cannot be unequivocally accepted as a valid instrument for distinguishing behaviorists from traditionalists.

The three raters utilizing this questionnaire to categorize the clinician samples found the best indicators of which theoretical orientation a therapist was allied with were the subject's self description, the rank ordering of the factors the subject felt were most important to consider when practicing psychotherapy, and the activities the subject felt were most important in his/her own brand of treatment.

It is unfortunate that in grouping clinicians for the current investigation, neither psychiatrists nor
psychiatric residents were found in the samples. Historically, psychiatrists have been more analytically and medically oriented than behaviorally oriented; it would have been interesting to examine the relationship between the degree to which they might be influenced by diagnoses and their training which is quite different from that of psychologists and clinical social workers. However, the psychiatrists and residents contacted for this study were unwilling to participate in this type of research. Apparently they were less interested in the results of scientific investigations into "therapeutic orientation" as discussed in the cover letter than the other mental health professionals. Perhaps, future researchers will be more successful in capturing this rich, and possibly unique subject pool.

Future Research

The present investigation concerned itself primarily with the relationship between a labeling effect and ones theoretical orientation and training. Further research in this area might include some improvements in the labeled stimuli--perhaps four comparable films, two boys (one disturbed and one normal) and two girls (one disturbed and one normal); and more attention paid to the process of subjects' attention to and integration of the diagnostic information. Perhaps a more formal method of
assessing whether subjects have attended to the diagnoses, understood their definitions, associated them with other information they have in storage, and utilized them when evaluating the stimuli, could be designed. Two other important questions should be: Are observers more apt to be influenced by a high prestige or relevant communicator than a low prestige or irrelevant communicator? (Temerlin, 1968) and What cues are observers using when rating behavior? (Blumberg, 1971). In addition, if two films are used in replication studies, it might be helpful to counterbalance the film showings for order effects.

This investigator is also concerned with how other samples might be affected by labeling. It would be interesting to replicate this study employing a physician sample, a nurse sample, a special educator sample, a parent sample, a non-parent sample, a child clinician sample, an adult clinician sample, and even "normal" and "disturbed" child samples. Other clinician variables which could be fruitfully examined in their relationship to labeling bias are: sex, empathy, warmth, authoritarianism, political views, years of mental health experience, and formal exposure to children. Other variables which tend to keep observers from making objective observations are listed on page 33 (Cline, 1964). These also might be examined in relation to a labeling effect.
It might be interesting, in addition, to develop other stimuli rating devices which might tap subjects' feelings of attraction or repulsion for the labeled children, their preferences for the treatment modality they might use for such children, and their prognoses for the children. A comparison of the labeling of children with the labeling of adults would be a socially relevant investigation, as would an examination into the nature of the labels—e.g., physical handicaps versus mental handicaps. Within the stimuli or the pre-film diagnostic information, factors of sex, race, socioeconomic status, and religion could also be manipulated and analyzed.

Certain pretherapy events such as the variables just mentioned as well as diagnosis can have a great deal of bearing on how the patient and the therapist relate to each other in their first and subsequent sessions. An analogue experiment along the lines of Gustin's (1969) study designed to tap a therapist's attitude toward the client after the therapist receives such pretherapy information might also be revealing.

Implications

While specific implications of findings from this investigation have already been noted, there are some general considerations which should be discussed. The current study, to the extent that it is an outgrowth of social
perspective investigations, examined the process of diagnostic labeling or mislabeling with respect to its biasing effect on observers. It was found that diagnoses did tend to influence and distort the perception of those inexperienced observers and some of the experienced clinicians in a position to evaluate the filmed children's behavior. In "social perspective" terms it becomes apparent that diagnostic labels employed injudiciously may lead to "devaluation" (see Wolfensberger, 1972) of the labeled individual. Becker (1963) sees the deviant in society as merely one to whom the diagnosis has been successfully applied. This investigation demonstrated that once a child is so labeled, it is possible that many of his/her behaviors and characteristics will be colored by that label.

Psychologists and other mental health workers in the past have not taken the social perspective research on diagnostic fallibility to heart. One reason is the methodological criticisms lodged against the research. The Braginskys (Braginsky & Braginsky, 1973) feel another major reason professionals and paraprofessionals are slow about applying the knowledge gained from the results of these studies in their work settings is that they are as biased as everyone else in our society. The Braginskys see the psychological classification system as an outgrowth and reflection of an entire conservative, middle-class political system within which these service providers
operate. "The examination of diagnostic labels historically, linguistically, and empirically makes it clear that these labels tell us nothing about the labeled but a good deal about the labelers and the society they serve" (Braginsky & Braginsky, 1973, p. 112).

A third reason psychologists have been slow to react to research such as Rosenhan's, Langer and Abelson's, and Temerlin's, into diagnostic labeling, is that they have been trained in the utilization and affixing of such labels; diagnostic techniques they employ are geared to yield such labels; plus administrative, government, health, and educational bureaucracies have been constructed to handle individuals who have been categorized and described with the traditional labels. Psychologists in applied settings are often quite critical of the current classification schema and are sometimes aware of the biasing effect it can have. However, until they have some viable modifications or alternative systems to fall back on, they are unwilling to heed research which suggests that the contemporary system be scrapped.

One psychologist who is offering a constructive alternative is Nicholas Hobbs. He sees psychological labeling or mislabeling as having the potential of being even more detrimental when used or imposed on "handicapped children." Trotter (1975) discussed Hobbs' recent compelling report to HEW on the effects of classification on children
in the APA Monitor. She wrote that:

classification can be used to sanction treatment of children in ways that no professional group defends and that labels . . . generate expectations that often work at cross purposes with the most enlightened efforts to help children (p. 5).

The current study points to the possibility of such a situation being created by the use of contemporary psychiatric diagnoses.

A massive screening program for 13 million Medicaid eligible children in the U.S. is currently being designed by HEW. This screening will include a battery of medical and psychological tests. Theoretically such screening of infants and young children could lead to corrective treatment of developmental and behavioral disabilities and the prevention of chronic, disabling conditions. However, some critics fear that such screenings can result in ambiguous and stigmatizing labels such as "minimal brain dysfunction," mental retardation, hyperkinesis, or "learning disabilities." These labels often contribute to exclusion to "special programs" or institutions rather than to treatment (Trotter, 1975). Too often it seems similar screening procedures, especially mass medical screenings, are implemented when treatment services are unavailable or service providers are not adequately trained to deliver it. In a stormy protest of HEW's screening plan, Catherine Jermany of the National Welfare Rights Organization stated: "There's
enough tracking in this country already, enough slipshod labeling. Poor people would rather be told that they're dying of cancer than that their kid is crazy" (Trotter, 1965, p. 23).

The question findings from this research raises is "Should diagnostic labels ever be used to categorize mental health/illness, and if so, what form should these labels take?" Virginia Satir (Note 7) stated that professionals too often use diagnoses as identifying tags leading to expectations and prejudices, rather than using them as descriptions. Many psychologists and psychiatrists have recently begun to use descriptive or discriminative systems they find more "useful" than the traditional illness categories of psychological dysfunction which are tied to the medical model. Menninger (1963) talks of "coping devices of everyday life" and "five orders of dysfunction." James and Jongward (1971) have simplified classification by labeling people either as "winners" or "losers" in life and describing the characteristics of each. Carkhuff (1969) and Egan (1972) use a five-point scale measuring overall psychological functioning, global helping ability, and competence in individual helping and human relations skills, as a discriminative tool.

Labels will always be demanded by those agencies responsible for maintaining statistics and by those responsible for funding "special" treatment programs;
therefore labeling will probably continue (Imhoff, 1973). However, if diagnoses focus more upon development of competencies and less upon pathologies, and as individual differences are accommodated more adequately within the setting of regular service systems, such as schools, children may not need to bear the burden of stigma producing labels (Clarizio & McCoy, 1976).

It is crucial to be aware of how easy it is to criticize, through one's research, the existing system of diagnostic classification. The more difficult, yet more important task, is the scrutinizing of this system with the purpose of developing constructive and humane alternatives to be used in the delivery of children's mental health services. It is this challenge that psychologists will hopefully rise to in future investigations.
REFERENCES


Cline, V., & Richards. Accuracy components in person perception scores and the scoring system as an artifact in investigations of the generality of judging ability. *Psychological Reports, 1963, 12,* 363-73.

Coleman, J. *Abnormal psychology and modern life.* Glenview, Ill.: Scott, Foresman, 1964.


Gallagher, J. The special education contract for mildly handicapped children. Exceptional Children, 1972, 38, 527-351


James, M., & Jongward, D. Born to win: Transactional analysis with Gestalt experiments. Reading, Mass.: Addison-Wesley, 1971.


Trotter, S. Labeling: It hurts more than it helps. APA Monitor, Jan., 1975, 6, 5-6.


REFERENCE NOTES


APPENDICES

APPENDIX A
Dear

We are involved in a research project focusing on a measure which will eventually be used to determine the therapeutic orientation of clinicians. Dr. has mentioned your name as a helpful person to contact for locating possible subjects. We are requesting your assistance in arranging to use the therapist trainees, psychology interns, psychiatric residents, or students you may supervise and the clinical staff members (therapists, analysts, etc.) you work with, as subjects.

Clifton Saper, an advanced clinical psychology graduate student at the Loyola Guidance Center, is the primary researcher, and Dr. James Johnson of Loyola University, Dr. Thomas Petzel of Loyola University, Dr. John Shack, Director of Research and Training at the Loyola Guidance Center, and Dr. Patricia Barger, Executive Director of the Loyola Guidance Center, comprise the committee supervising this enterprise.

We hope to be able to send out our Clinician Questionnaire to as many therapists as possible. We are defining therapist as any professional staff member or trainee who has at least one year of experience working directly with patients or clients in a clinic, hospital, special school, or private practice. The questionnaire should take no longer than 15 minutes to complete and can be filled out at the subjects' convenience. The design of the research demands the confidentiality of the subjects, so the therapists/subjects can participate without feeling their clinical skills are in any way subject to identification or evaluation.

The way you can help us in phase I of our research is to send me the list of names, work addresses, and phone numbers of those therapists (interns, residents, trainees, or professional staff members) you work with who might be willing to fill out our Clinician Questionnaire assessing one's therapeutic orientation. (Use the enclosed stamped envelope) With continued research on such a measure, we are hoping it may develop into a valuable tool for analyzing clinician variables and how they are related to therapeutic outcome. If you would like to give us the list of names over the phone rather than through the mail, that would be fine with us. You can call our primary researcher, Clifton Saper, at the Loyola Guidance Center at 274-5305 (or 5306).

After we have a large enough sample of therapists who have completed questionnaires, we will recontact some of them to request their participation in Phase II of our research which involves viewing some films. I very much appreciate your help and will be in touch with you on Jan. 13 or 14 to answer any questions you may have. Thank you very much for your valuable time and assistance.

Sincerely,

Clifton Saper
For the Research Committee
APPENDIX B
**CLINICIAN QUESTIONNAIRE**

Name: ______________________________  Sex: ___  Age: ______

Work Address: ____________________________________________________  Phone: ________

Professional Status:  (circle one)
1. Psychiatrist
2. Psychiatric Resident, year ______
3. Clinical Psychologist
4. Psychology Intern, year ______
5. Psychologist (M.A, Social, Developmental, Counseling, etc.)
6. Psychiatric Social Worker
7. Psychiatric Nurse
8. Social Work trainee
9. Psychoanalyst
10. Psychoanalyst-in-training
11. Special Educator
12. Trainee in Special Ed.
13. Other

Years of Clinical Experience: ________  Years of Clinical Experience with Children: ________

If you have completed your clinical training, at which facility did you receive the majority of this training? What was the major theoretical orientation of that program?

Briefly describe, as specifically as possible, your basic theoretical orientation, techniques, and goals in therapy:

```
(Use back of Sheet if necessary)
```

<table>
<thead>
<tr>
<th>1. How would you rate your adherence to a medical model of abnormal behavior (circle one)?</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-medical</td>
</tr>
<tr>
<td>model</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>2. How would you rate your therapeutic orientation on a psychoanalytic (ego psychology) --non-analytic continuum?</th>
</tr>
</thead>
<tbody>
<tr>
<td>non-analytic</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>3. How would you rate your therapeutic orientation on a behavioral -- global personality continuum?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Behavioral</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. Do you feel that mental illness is something ascribed to people as a function of definitions given certain types of acts by society or something ascribed to persons who are ill?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Societal</td>
</tr>
<tr>
<td>context</td>
</tr>
<tr>
<td>1</td>
</tr>
</tbody>
</table>
5. Abnormal behavior should be thought of as the product of an illness, a compensatory reaction to an organic defect or a combination of these?

<table>
<thead>
<tr>
<th>Disagree</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

6. The uncovering and resolution of emotional problems which are causing the patient's symptoms should be the main goal of therapy. Curing the symptom itself rarely solves the real problem. Disagree Agree

<table>
<thead>
<tr>
<th>Disagree</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

7. The most important goal of therapy should be the patient's learning of new skills which can be used in stressful situations.

<table>
<thead>
<tr>
<th>Disagree</th>
<th>Agree</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
</tr>
</tbody>
</table>

8. Do you feel that the examination of childhood experience is essential to effective adult psychotherapy? Not important Essential

<table>
<thead>
<tr>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
</table>

9. Do you feel that the use of official A.P.A. psychiatric diagnoses is helpful for therapists?

<table>
<thead>
<tr>
<th>A hindrance</th>
<th>Very helpful</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

10. Do you feel that problematic behavior produced by severe emotional disturbance or psychoses can be eliminated or changed with specific short-term therapeutic techniques or should treatment take the form of a long-term relationship?

<table>
<thead>
<tr>
<th>short-term therapeutic technique</th>
<th>long-term relationship</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

11. Do you feel that emotional disturbance in children is a medical problem best treated in a hospital setting, or is a problem in living best treated in a special school setting?

<table>
<thead>
<tr>
<th>hospital setting</th>
<th>special school setting</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
</tbody>
</table>

Which of these factors do you feel are most important to consider when therapy is done (Pick three and rank order them)?

- organic defects
- sources of reinforcement in the individual's social environment
- unconscious psychological processes
- the description and frequencies of specific observable behaviors
- the global personality description
- intrapsychic conflicts
- insight
- learning patterns
- catharsis
- defense mechanisms

Rank order, by importance to your own brand of therapy, these 13 activities.

- concentrating on childhood relationships and psychosexual conflicts
- actively interpreting
- being empathic
- keeping base rates and charts of behavioral frequencies
- stimulus control and consequence analysis
- dream interpretation
- intrusive physical contact
- encourage free association
- interpret transference
- suggesting action programs and homework assignments
- breaking problem into component parts and working on them
- using role playing
- using relaxation techniques or systematic desensitization

Check this box if you would like to be informed of the results of this research [ ]
We are conducting research on the theoretical orientations of therapists. Your name was suggested to us by , who said you would probably be willing to assist us by filling out the enclosed questionnaire.

The Clinician Questionnaire should take you less than 15 minutes to complete. We are defining therapist as anyone (trainee, intern, resident, professional, or trained paraprofessional) with at least one year of experience working directly with disturbed adults or children in a hospital, clinic, special school, or private practice. This definition most probably includes you.

The research committee who contacted includes: Clifton Saper, M.A. primary investigator and Administrative Assistant of the Loyola University Guidance Center; James Johnson, Ph.D.; Loyola U.; Thomas Patzel, Ph.D., Loyola U., John Shack, Ph.D., Director of Research and Training at the Loyola Guidance Center; Patricia Barger; Ph. D., Executive Director of the Loyola Guidance Center. We all thank you very much for your assistance in this project and greatly appreciate the time you are devoting to it.

Your name, phone number, and address are only being used to sort the data, to possibly recontact you for the second phase of our research, and to provide you with the results of our study (you can be informed of the results by checking the box on the bottom of the questionnaire). As soon as possible your name will be changed to a number. Every measure will be taken to insure your confidentiality and privacy. In absolutely no circumstance, will the data you have provided be used in evaluating your clinical skills or training, and it will not be made available to anyone besides the primary investigator.

We hope that you will complete this questionnaire at your earliest possible convenience (Please use the enclosed stamped envelope to return it to us). We hope to have all data in by early February. Clifton Saper may be calling you next week to make certain you have received this material and answer questions you may have. Feel free to call us if you have any concerns about this research.

When all the data has been gathered, we will begin Phase II of our project. This may involve recontacting you to participate in another task. This task consists of viewing two short and enjoyable films and answering some questions about them. These films will be shown at your own facility or office at your convenience.

Thank you again for your assistance in this project.

Sincerely,

For the Research Committee
Loyola Guidance Ctr.
1043 W. Loyola Ave.
Chicago 60626
274-5305 (or 5306)
Description of Children in Films by a Professional Developmental Psychologist

Girl

The girl was an attractive child who appears uncertain and somewhat tense, especially in relation to other children. She appeared interested in their activities but did not participate (or is not accepted?). In reaction, she releases energy or tension by solo performances of hand springs, attempting to stand on her hands, etc. Such activity may also serve to gain approval from adults. However, a problem in relating to others seemed supported by lunch behavior—very concentrated in food with no attention to others present. Her coordination and ability to concentrate appear good (performance on tasks and gymnastics). On tasks, she was a quick worker but willing to persist even when having some problems (block construction). She is probably quite bright. She seemed to be more relaxed in relation to things than to people.

Boy

On playground, he appears active and eager to participate with other children, but he may lack the necessary skills. When he fails to get ride on cart, he runs off yelling and then joins adults. While eating, he shows affectionate relationship with adult (touch on arm). In tasks, he seems to lack skill or motivation—especially the latter. He appears bored and dejected. I expect he likes to be active, likes attention, and is minimally interested in school.
II

Description of Children in Films by an English Teacher

Girl

Though not an unfriendly person, this girl sometimes was very willing to engage in activities with others. At other times she was unpleasant and refused to participate. She was especially pleased when she became the center of attention. Like many kids, she would make humorous faces or play with her food. She was also on the competitive side when she felt she could win, but would run away from situations she felt unsure of. She was also non-verbal at times and her play was erratic.

Boy

This boy seemed to be very verbal and active at all times. He easily joined in the activities with the other children and was only upset once, and he went and talked to other people. His eating was fairly good, yet he opened his mouth and clowned around a bit. His concentration was good when he was given things to do. When he couldn't figure out what to do, he took the easy way out.
III

Description of Children in Films by an Untrained Undergraduate Psychology Student

Girl

The girl in the film looked approximately six or seven years old. She had an intelligent face and was of average size. Her motoric ability was well-developed; she walked, ran, and did handstands on the lawn. The only unusual thing I noticed is that the girl did not experience any close physical contact with anyone. She didn't hug or kiss the adult or any of the other children. Even in crowd scenes she tended to be on the outside and aloof.

Boy

The boy in the film appeared quite normal. He seemed friendly and open. He played well with others and had well-developed motor reflexes. I believe, however, that the boy was deaf. It appeared in one scene that the woman watching him eat spoke to him without eliciting any response. Also, when people spoke to him they tended to use their hands more expressively. At one point an adult pulled him away from one of the go-carts after the little boy failed to respond to him.
Please go through the 24 items on this sheet and decide which could apply to the child starring in the film you just saw. If you think the item could apply, put an "X" on the blank in front of the item. Next go back over only those items you checked and decide to what degree you think the statement applies to the child in the film. The number "1" means the statement is slightly accurate in describing the child or his behavior or what it might mean. The number "2" means the statement is somewhat accurate in describing the child. The number "3" means the statement is quite accurate in describing the child.

The child in the film I just saw was (CIRCLE ONE) the BOY or the GIRL.

Key

Positive 1. This child is actively eager to participate in games with other children.

Positive 2. The child has an excellent ability to concentrate and sit still long enough to do a problem even if it is frustrating.

Negative 3. The child seems worried, moody, and pouty, rarely initiating any interactions with others. He or she also shows some irritability and possessiveness.

Negative 4. Child’s play behavior, when alone, seems primitive, sporadic, and well below that expected of his or her age.

Positive 5. The child plays with the other children, but is not one of them.

Positive 6. The child is energetic and has excellent fine and gross motor and visual skills.

Negative 7. The child’s eating habits are quite poor.

Negative 8. The child responds to relatively instructional situations with some passive-aggressiveness, hyperactivity, lack of patience, boredom, and difficulty in attending to and completing the assigned tasks.

Positive 9. The child is happy, in good spirits, and responds in a very natural way.

Negative 10. The child’s behavior suggests a relatively schizophrenic or autistic adjustment pattern with a great deal of anxiety about his or her own physical and emotional safety. The child seems un related to the things and people around him or her.

Positive 11. Like many kids, this child makes funny faces and plays with his or her food.

Positive 12. The child has good eating habits.

Positive 13. The child is innovative and creative, and his or her intellectual capacities seem within average limits.
Negative 14. The tasks the child performed are appropriate for a much younger child, yet he or she still has difficulty with fine motor coordination tasks.  
Positive 15. The child is competent, bright, self motivated, and pleased with his or her efforts and successes.  
Negative 16. There seems to be possible psychogenic autism and/or schizophrenic adjustment in this case.  
Negative 17. The child appears to be aimlessly throwing himself around while playing rather than being exuberant.  
Negative 18. The child’s behavior appears age inappropriate and developmentally retarded. He or she has a very short attention span, minimal verbalizations, with productivity and creativity being below age level.  
Positive 19. The child is affectionate with adults and generally responds happily and playfully to their teasing.  
Positive 20. The child seems to demonstrate a decisive and quick plan of action in structured, task oriented situations. He or she is a quick worker and is willing to persist even when having some difficulty.  
Negative 21. The child, during one segment of the film, appeared to be hallucinating or at least losing contact with events and circumstances around him or her. There is also a tendency toward postural stereotyping such as rocking and hand clasping.  
Positive 22. The child talks and relates adequately with others. He or she listens to and understands directions well, cooperating willingly with teachers and peers.  
Negative 23. The child responds more frequently to adults than to his or her peers suggesting that he or she may be working through significant dependency issues with his or her parents. The child always seems to be controlled by the adults around.  
Negative 24. The child does not like to be touched by adults or peers. He or she is unwilling to participate in activities and generally appears tense, inhibited, and uncertain in relation to other children. There is some difficulty in interpersonal relations and he or she is more relaxed in relation to things than to people.
APPENDIX F
<table>
<thead>
<tr>
<th></th>
<th>very</th>
<th>moderately</th>
<th>slightly</th>
<th>moderately</th>
<th>very</th>
</tr>
</thead>
<tbody>
<tr>
<td>active</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>extroverted</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>sociable</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>cruel</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>conscienceless</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>happy</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>dull minded</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>loving</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>trusting</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>quick</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>curious</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>optimistic</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>warm</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>responsive</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>adventurous</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>soft-hearted</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>colorful</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>outgoing</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>irritable</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>real</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
</tr>
</tbody>
</table>

- inactive
- introverted
- unsociable
- kind
- conscientious
- depressed
- intelligent
- not loving
- distrustig
- slow
- uninquiring
- pessimistic
- cold
- aloof
- timid
- hard hearted
- colorless
- self-centered
- easy going
- unreal
<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>prone to anger</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>not prone to anger</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>meaningless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>meaningful</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>interesting</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>boring</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>confident</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>feels inadequate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>formed</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>formless</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>noisy</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>quiet</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Film of boy masculine</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>feminine</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Film of girl</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>likes school</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>dislikes school</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>poor memory</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>good memory</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>excitable</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>calm</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>interested</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>bored</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>disobedient</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>obedient</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>truthful</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>lying</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tense</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>relaxed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>emotional</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>selfcontained</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>strong willed</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>weak willed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>independent</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>dependent</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>attention avoiding</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>attention seeking</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>irresponsible</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>responsible</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

very | moderately | slightly | slightly | moderately | very
<table>
<thead>
<tr>
<th></th>
<th>very</th>
<th>moderately</th>
<th>slightly</th>
<th>slightly</th>
<th>moderately</th>
<th>very</th>
</tr>
</thead>
<tbody>
<tr>
<td>not helping</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>helping</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>obstructive</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>cooperative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>effective</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ineffective</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>adjusted</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>maladjusted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>friendly</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>not friendly</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>happy</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>sad</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>leader</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>follower</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>always on the go</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>not active</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>never seems to tire</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>tires easily</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>outdoor type</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>indoor type</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>emotionally healthy</td>
<td>6</td>
<td>5</td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>emotionally disturbed</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Please complete this form as if you had been observing the child in the film at home and in school over a long period of time. Indicate which of the following might constitute problems as far as this child is concerned. If you guess that an item would not constitute a problem, circle zero; if you guess that an item would constitute a mild problem, circle the one; if you guess that an item would constitute a severe problem, circle the two. Please use your imagination to predict or extrapolate answers from the child's filmed behavior and complete every item.

Circle one: Completed as if having observed the BOY GIRL.

1. Thumb-sucking
2. Restlessness, inability to sit still
3. Attention-seeking, "show-off" behavior
4. Skin allergy
5. Doesn't know how to have fun; behaves like a little adult.
6. Self-consciousness; easily embarrassed
7. Headaches
8. Disruptiveness; tendency to annoy and bother others
9. Feelings of inferiority
10. Dizziness, vertigo
11. Boisterousness, rowdiness
12. Crying over minor annoyances and hurts
13. Preoccupation; "in a world of his own"
14. Shyness, bashfulness
15. Social withdrawal, preference for solitary activities
16. Dislike for school
17. Jealousy over attention paid other children
18. Prefers to play with younger children
19. Short attention span
20. Bedwetting
21. Inattention to what others say
22. Easily flustered and confused
23. Lack of interest in environment, generally "bored" attitude
24. Fighting
25. Nausea, vomiting
26. Temper tantrums
27. Restlessness, irritability
28. Truancy from school
29. Hypersensitivity; feelings easily hurt
30. Laziness in school and performance of other tasks
31. Anxiety, chronic general fearfulness
32. Irresponsibility, undependability
33. Lack of self confidence
34. Excessive daydreaming
35. Tension, inability to relax
36. Disobedience, difficulty in disciplinary control
37. Depression, chronic sadness
38. Uncooperativeness in group situations
39. Aloofness, social reserve
40. Passivity, suggestibility; easily led by others
0 1 2 41. Clumsiness, awkwardness, poor muscular coordination
0 1 2 42. Stuttering
0 1 2 43. Hyperactivity, "always on the go"
0 1 2 44. Distractibility
0 1 2 45. Destructiveness in regard to his or her own and/or others' property
0 1 2 46. Negativism, tendency to do the opposite of what is requested
0 1 2 47. Impertinence, sauciness
0 1 2 48. Sluggishness, lethargy
0 1 2 49. Drowsiness
0 1 2 50. Profane language
0 1 2 51. Prefers to play with older children
0 1 2 52. Nervousness, jitteriness; easily startled
0 1 2 53. Irritability; hot tempered, easily aroused to anger
0 1 2 54. Stomach aches, abdominal pain
0 1 2 55. Specific fears, e.g. of dogs, of the dark, of riding in or on a vehicle.
The child in the short film you are about to see is a normal six year old girl who was filmed while visiting a special school at which her father is an administrator. The day school for emotionally disturbed children is affiliated with a local public school in Chicago. Your task is to carefully watch the short film which follows that focuses on this child. Do not turn to the next page in this packet until you are told to when the film is over.

Once again, we thank you very much for participating in this research and giving us an hour of your valuable time.
The child in the short film you are about to see is a normal five and a half year old boy who attends a parochial school in Chicago. One of his sisters is a student volunteer at a special school in the area. He was filmed on one of his vacation days while visiting her as a guest at this school. Your task is to carefully watch the film which follows that focuses on this child. Do not turn to the next page in this packet until you are told to when the film is over. Once again, we thank you for giving us an hour of your valuable time to participate in this study.
The child in the movie you will be seeing next is a six year old girl who has been excluded from the Chicago Public Schools and attends a special school for emotionally disturbed children in the area. She is being filmed at this school. She has been given the combined diagnosis by a psychologist working for the Board of Education and her therapist of mental retardation and severe emotional disturbance produced by a symbiotic psychosis. Your task is to carefully watch the short film which follows that focuses on this child. Do not turn to the next page of this packet until you are told to when the film is over.
The child in the film you will be viewing next is a five and a half year old boy who was recently tested by the Bureau of Child Study in Chicago where he was given the following diagnosis by the psychologist and psychiatrist who saw him: severe childhood schizophrenia involving pre-psychotic symbiotic ties; mild mental retardation; and epilepsy. He is being filmed at the special school for emotionally disturbed children which he attends. Your task is to watch carefully the short movie which follows that focuses on this child. Do not turn to the next page of this packet until you are told to when the film is over.
The dissertation submitted by Clifton J. Saper has been read and approved by the following Committee:

Dr. James Johnson, Chairman
Associate Professor, Psychology, Loyola

Dr. Patricia Barger
Professor, Psychology, Loyola

Dr. Thomas Petzel
Associate Professor, Psychology, Loyola

The final copies have been examined by the director of the dissertation and the signature which appears below verifies the fact that any necessary changes have been incorporated and that the dissertation is now given final approval by the Committee with reference to content and form.

The dissertation is therefore accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Psychology.

Date: 2-17-76

Director's Signature