DOCTORAL DISSERTATION SERIES

PUBLICATION: 5037

AUTHOR: Victor Goertzel, Ph. D., 1953
University of Michigan

TITLE: SHIFTS IN PERSONALITY IN THE RORSCHACH TEST AND IN PSYCHOTHERAPY

University Microfilms, Ann Arbor, Michigan, 1953

SHIFTS IN PERSONALITY IN THE RORSCHACH TEST AND IN PSYCHOTHERAPY

by
Victor Goertzel

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy in the
University of Michigan
1953

Committee in charge:

Associate Professor Max L. Hutt, Chairman
Assistant Professor Daniel R. Miller
Assistant Professor George A. Satter
Assistant Professor Herbert T. Schmale
Doctor Thornton W. Zeigler

SHIFTS IN PERSONALITY IN THE RORSCHACH TEST AND IN PSYCHOTHERAPY

(Abstract)

By Victor Goertzel

This study is concerned with predicting therapeutically induced changes in behavior from an evaluation of variables especially derived from Rorschach's test. Therapy is regarded primarily as a condition providing for the induction and observation of behavioral and perceptual change in the patient.

The primary experimental condition consisted of the re-administration of the Rorschach test with altered instructions, requesting the subjects to modify their performance in the direction of more Factor. A second experimental condition consisted of the administration of an "ink blot" test designed to establish the full extent of the subjects' capacity to utilize this significant Rorschach variable, human movement.

The subjects were 40 young adult psychoneurotics. Following acceptance for treatment at a public mental health clinic they were treated under the circumstances usually obtaining there. Evaluations of the subjects' severity of illness, prognosis and desirability were obtained from the intake psychiatrists. Approximately six months after the treatment had begun, the therapists evaluated the changes resulting from [illegible] of therapy. In addition, psychiatric evaluations of therapeutic change were obtained from the psychiatric consultant.

A supplementary study of the relationship between manifest and latent dependency was undertaken to assist in evaluating some of the results of the study. The findings with respect to shift in the test situation were compared with the findings of a previous study with college students as subjects.

Within the limits of the study, eight principal findings are presented. (1) The human movement factor in the Rorschach is unstable. (2) There is no

[illegible] (2) The [illegible] (4) [illegible] (5) [illegible] (6) The [illegible] (7) [illegible] (8) The [illegible]

[illegible]

Three general conclusions were drawn from the results obtained in the study. (1) Women [illegible] (2) The capacity to shift in the [illegible] (3) Shift in the [illegible]

The author's [illegible] interpreting the study as indicating definitely that [illegible] The limitations on the extent to which the findings and conclusions of the study may be generalized are [illegible]. Suggestions are given for further research which would lead to the development of a more appropriate theoretical framework.

ACKNOWLEDGEMENTS

I am very indebted to Dr. Louis Koren, psychiatrist-in-chief at the Wayne County Mental Health Clinic for his facilitation in many ways of the study at the Wayne County Mental Health Clinic.

I owe a deep debt of gratitude to my advisor, Dr. Max L. Hutt, for his constant encouragement during the planning and execution of the study in its every phase and for his very helpful criticism. Valuable suggestions were obtained from other members of my doctoral committee: Dr. Daniel R. Miller, Dr. George A. Satter, Dr. Herbert T. Schmale and Dr. Thornton W. Zeigler.

This study would not have been possible without the cooperation of a great many people: the entire staff of the Wayne County Mental Health Clinic; the psychologists who did the Rorschach testing, particularly James M. Joyce, Dr. Richard N. Filer, and Spencer Sterne, who did the bulk of the testing; and to the patients who served as subjects.

Finally, for her patience and warm encouragement, my deepest affections go to my wife Mildred, to whom this work is dedicated.

TABLE OF CONTENTS

		Page
LIST OF TABLES.		v
LIST OF FIGURES		vii
Chapter		
I.	STATEMENT OF PROBLEM	3
II.	HYPOTHESES.	12
III.	EXPERIMENTAL DESIGN	17
	Subjects	17
	Procedure.	19
	Examiners.	21
	Scoring	21
IV.	ANALYSIS OF DATA	23
	Initial Rorschach Factors in Control and Experimental Groups	24
	M Set Change in Control and Experimental Groups.	25
	Change in M:C Ratio of Control and Experimental Groups	30
	M:C Ratio in Normal and Neurotic Groups.	32
	Relationship Between Initial M and M Set Change	36
	Stability Scores	37
	Exploratory Study of Initial High M Patients	45
	Psychiatrists' Initial Ratings and M Set Change,	47

Chapter Page

Evaluation of Therapeutic Change. . . . 54

Refinement of Measures of M Set Change. . 63

The M Limits Test. 65

Differences Between Therapists' Average Ratings of Therapeutic Change. 67

Consultant's Evaluation of Therapeutic Change 69

Rorschach Neurotic Signs and Therapeutic Change 73

Dependency Needs and Shift in the Rorschach 76

V. SUMMARY OF FINDINGS AND DISCUSSION 81

VI. GENERAL SUMMARY AND CONCLUSIONS 93

APPENDIXES

I. 102

II. 103

III. 105

IV. 110

BIBLIOGRAPHY 112

LIST OF TABLES

Table		Page
1.	Comparison of Initial Test of Control Group with Initial Test of Experimental Group for Selected Rorschach Variables. . .	26
2.	Comparison of Human Movement Response on Initial Test with Repeat Test for Control and Experimental Groups.	29
3.	Changes in Experience Type for Control and Experimental Groups.	29
4.	The Number of Subjects in Each Group for Both Original and Repeat Test with the Specified Direction of the E.B. (A Comparison of Milton's Data with the Data of the Present Study)	33
5.	Relationship Between Membership in a Group and Shift in Direction of E.B. . . .	34
6.	Relationship Between Membership in Normal Control and Neurotic Control Group and Shift in Direction of E.B. . . .	34
7.	Relationship Between Membership in Normal Experimental and Neurotic Experimental Group and Shift in Direction of E.B.	35
8.	Milton's Stability Score Factors with Chi Square and P Values for Neurotic Experimental Group Dichotomized in Most Shift (N=13) and Least Shift (N=13) Sub Groups.	41
9.	Goertzel's Stability Score Factors with Chi Square and P Values for Neurotic Experimental Group Dichotomized into Most Shift and Least Shift Sub Groups	41

Table		Page
10.	Comparison of Stability Scores of Sub Groups of Experimental Group Dichotomized as to Most and Least Change in M Set Retest.	41
11.	Comparison of Two Groups of High Initial M Subjects	49
12.	Relationship Between Score for Psychiatrists' Over All Initial Rating and M Set Change	51
13.	Relationship between Score for Psychiatrists' Ratings of Severity of Initial Illness and M Set Change.	52
14.	Relationship Between Psychiatrists' Rating of Compulsivity and M Set Change	53
15.	Chi Square Test for Significance of Differences Between Average Therapists' Ratings of Therapeutic Change	68
16.	Neurotic Scores (6 of 8 Miale and Harrower-Erickson Neurotic Signs) and Scores for Therapists' Ratings of Therapeutic Change	74
17.	Relationship Between Number of Neurotic Signs and Therapeutic Change	75
18.	Relationship Between Number of Neurotic Signs and M Set Shift	76

LIST OF FIGURES

Figure		Page
1.	Initial Rating Form Used by Intake Psychiatrists with Weights Given to Each Item.	50
2.	Rating of Therapeutic Change Form Used by Therapists with Weights Given to Each Item.	56
3.	Frequency of Response to Each Item on the Therapists' Rating of Therapeutic Change Form.	57

CHAPTER 1

STATEMENT OF PROBLEM

There is a prevailing interest in psychology in the relationship between variables in the personality and the prediction of therapeutic change. This study is concerned with predicting therapeutically induced changes in behavior from variables especially derived from Rorschach's test. For the purpose of this study, therapy is regarded as a prolonged interpersonal relationship which provides a maximal condition for the induction and observation of behavioral and perceptual change in the patient.

The person who, in a test situation, is able to restructure his perception of a visual stimulus should, in the therapeutic situation, have a greater capacity toward restructuring his perception of his "self" and the important people in his life than does the person not so endowed. A person so endowed should be able, on request, to see more human movement responses on a Rorschach retest than he saw during a previous Rorschach given in a conventional manner. The perception of human movement (M) in the Rorschach should be an indication of the capacity for developing relationships to people and of the nature of such relationships. The amount and kind of M responses should

have predictive value for therapeutic change. An increment in prediction should be obtained through use of a measure of the ability to increase the perception of M. Possible explanations for an increase in M are discussed later. A further increment in prediction should result from a measure of latent capacity for M responses not revealed in the original test or in the M-set retest. A systematic testing of the limits for M following the retest should be an indication of this latent capacity.(11) The core of this research is contained in the preceding premises.

Considerable work has been done on the effect of varied experimental sets upon Rorschach performance by Hutt, Gibby, Milton and Pottharst (12, 9, 19, 22) and by Lord, and Carp and Shauzin (16, 7). Of particular interest was Milton's finding, (19) that college students were able almost to double their average number of human movement responses without any loss in the "quality" of the perception, simply by being asked on the test to tell everything they saw and in addition to find as many human movement responses as they could. This finding is in sharp contrast to Rorschach's own conception concerning human movement as a "deep", largely unconsciously determined variable. One of the problems of this investigation was to test the extent to which psychoneurotics would confirm Rorschach's expectations.

This increase in human movement responses (M) upon request in the direction asked for can be conceived as a

capacity to shift and may be developed as a means whereby we can assess an individual's limits and capacity for change in a test situation. This capacity to shift may be related to adaptability in life situations and may be developed as an important differentiating criterion of mental health.

In the present investigation the major operations of the Milton study on college sophomores are applied to a psychoneurotic population to test the basic hypothesis that capacity to shift is related to behavioral change. An additional measure of capacity to shift is obtained through a testing the limits for M procedure. This increase in M perception is related to therapeutic responsiveness in an attempt to get an increment in the prediction.

Successful therapy requires a change in the patient's perception of himself and of the significant persons in his life. The common phrase, "the patient has come to see things differently" is more than a loose descriptive analogy. A study by Jonietz (13) shows 1) that change and reorganization of visual perception occur during client centered therapy, and tend not to occur in a control group; 2) the change which occurs is a generalized one in that it affects the perception of an unstructured stimulus which was not a topic considered in therapy; 3) further that the reorganization initiated by therapy doesn't stop with the completion of the therapeutic interviews; 4) the degree of change and reorganization in visual perception is a real indicator of therapeutic change as it is correlated with the depth or

satisfactoriness of therapy as judged by other criteria; and 5) the types of perceptual change which occur are among those ordinarily associated with improved adjustment according to subjective clinical judgment.

To the extent that personality reorganization is accompanied by perceptual reorganization, a clinical measure of capacity for perceptual reorganization would be a valid index of capacity for personality reorganization. The Rorschach technique which studies the subject's perception of ten standardized but ambiguous visual fields should lend itself to a study of the relationship between flexibility in visual perception and flexibility in therapy.

There are several broad questions which have been selected for investigation in this study. The first major question is: Is there any relationship between capacity to shift in variables in the Rorschach and therapeutic change? Subsumed under this question is another: Is there any relationship between capacity to shift on the Rorschach and behavioral characteristics of a patient, such as degree of illness, prognosis and compulsivity as judged clinically?

The second major question is: To what extent may personality resources not evident in the standard Rorschach procedure be uncovered in a retest under a special set and by a systematic testing the limits procedure? Subsumed under this question: Does the retest under the experimentally induced change in set increase the score on a Rorschach variable proportional to the amount of that variable produced

in the original test? Is the person who sees more M in the original test the one who has a greater potential capacity for the perception of M and therefore the one who will increase his M production under the set retest? May some persons who perceive no M on the original test actually be suppressing or repressing or for some other reason produce no M responses, and still have a capacity for making M responses and do so under the set induced by the conditions of the experimental retest? Does the testing the limits increase the score on a particular variable proportional to the amount of that variable revealed in the original Rorschach, or in the M set retest?

The answers to the above questions cannot be approached until we have some idea as to the answer to our third major question: To what extent are the scores on Rorschach test variables modifiable by psychoneurotic individuals, when the subject is given specific brief instructions to attempt to alter these scores. Under this first broad question are significant specific questions, such as: Is the change in a given variable correlated with any type of alteration in other variables, separately or in combination?

In this study we have worked with the Rorschach test and in particular with the human movement variable. The general problem of the relationship between capacity to shift in a test situation and ability to change in therapy could have been approached in many ways. We chose to work with the

Rorschach. The Rorschach appears to be the most widely used instrument in clinical psychology. Burton (6) finds that the Rorschach is the only test used in every one of the cases in a text containing reports submitted by 45 clinicians. The fact of its widespread use in clinical practice and in research studies attests to its usefulness. Our position is that unless there is some justification for new approaches in the way of greater or more specific information yielded, or greater economy, it is desirable to increase the utility and the theoretical understanding of techniques about which a great deal of normative and other data are already available and which have won a place for themselves in the psychologist's armamentarium.

Even a superficial review of the extensive research which has used the Rorschach method would be a major task in itself and will not be attempted here. The most comprehensive bibliographical coverage up to 1948, is given by Bell.(4) His chapter on the Rorschach technique includes a 798 item bibliography. An indication of the scope of research activity using the Rorschach now going on, is the fact that in her 1951 summary of current problems in Rorschach theory and technique, Hertz (10) makes reference to 193 studies, published and unpublished, mostly recent ones.

The human movement variable was selected for study because it has generally been believed that it reveals deep unconscious aspects of the personality and that it is

extremely resistant to change. Yet, Milton has shown that a sample of college students can almost double their production of M when they are simply asked to do so. The question of whether neurotics are able to do the same seems significant for personality theory, for projective test theory and for practice.

There is evidence that the presence of a number of M's in a record is a favorable prognostic sign, at least in schizophrenics, although the question of just what M represents is still not definitely determined. Piotrowski's (21) improved insulin patients had five times as many M responses as his unimproved patients. Stotsky (25) finds more M in his remitted group and includes M in his prognostic index.

Rorschach believed that "In normals, the number of M responses rises in proportion to the productivity of the intelligence, the wealth of associations, the capacity to form new associative patterns (23; p. 26)." Klopfer (14) believes that the number of M provides a basis for judging the subjects' independence and self-sufficiency and that it is somehow related to self-insight. Most writers, including Beck (4), Klopfer (14) and Rorschach (23), agree that it is a measure of current intellectual functioning. After reviewing the pertinent literature, Barrell (1) concludes that none of the studies on M as a measure of intellectual functions is conclusive.

Hutt believes that the human movement response represents a stablizing influence against the acting out or

direct expression of impulses.[1] He regards it as a functional barrier. Hutt's view is congruent with Rorschach's statement in his analysis of a test protocol, that "the individual who is influenced by kinesthetic factors in the test is stable in his general motility; the energetic person is influenced little by a sense of movement in the figures". (23; p. 25)

The human movement response is one of the first things the clinician turns to in evaluating a patient through a Rorschach test. The clinician considers the number of M responses, the relation of M to other variables, particularly to the color determined responses and the number of responses to the whole card. Some clinicians consider the location of the M responses, the form quality, the flexor or extensor type of movement and the sex of the figure. From the type of M the clinician gets hunches as to the patients relationship to male and female figures, how he handles aggression and other aspects of his interpersonal relationships. Writers on Rorschach interpretation from Hermann Rorschach to current text book authors, Beck, Klopfer, Rappaport, Bell, all make much of the M response. Beck regards "Rorschach's penetration to the essence of the movement response as fantasy activity" as his greatest single achievement. (4; p. 22) Beck starts his interpretation of the fundamental structure of the personality by considering the M and the C and the F+.

---

[1] Lecture notes, 1949

These factors, he contends, are the skeleton of the psyche. All the other test factors are essential but seem only to add flesh to the skeleton. There is a need to establish the full extent of a subject's capacity to utilize significant Rorschach determinants. The testing the limits for M following an M set Rorschach test, is an attempt to estimate this capacity for one most significant determinant.

CHAPTER 2

HYPOTHESES

Our fundamental hypotheses is that the capacity of the neurotic individual to shift his M score on the Rorschach on the basis of an experimentally induced set (as defined below) is significantly related to capacity for behavioral change.

The capacity to shift in M is measured by:

1. The amount of increase in M in the second test given under the experimental conditions compared to the amount of M in the first test.
2. The amount of increase in M in the M limits test.
3. The amount of increase in M in the experimental retest plus the M limits test.

The control group affords an opportunity for assessing the effects of a second administration of the Rorschach to a neurotic group within a short period of time. The control group is used to assess the changes which are due to the experimental condition in that the retest changes must be taken into consideration in evaluating the effect of the experimental set.

The capacity for behavioral change is measured by the therapists' ratings of therapeutic improvement.

The fundamental hypothesis that the capacity of the neurotic individual to shift his M score on the Rorschach is significantly related to capacity for behavioral change, then, will be tested in a series of specific hypotheses, thus:

1. The amount of increase in M in the second Rorschach given under the experimental conditions will be significantly related to therapists' ratings of therapeutic improvement.
2. The amount of increase in M in the M limits test will be significantly related to therapists' ratings of therapeutic improvement.
3. The amount of increase in M in the experimental retest plus the M limits test will be significantly related to therapists' rating of therapeutic improvement.

Our second major hypotheses is that the experimental neurotic group as a whole will shift its set on the Rorschach in the M factor significantly less than the experimental normal group as a whole.

Both of our major hypotheses emerge directly from our conception of the nature of the capacity to shift. This

conception[1] is that the healthy organism is sufficiently sensitive to variations in situations which confront him, and sufficiently adaptive to these variations, to be able to modify his responses so as to deal effectively with them. The healthy organism is highly differentiated and adaptive. Its defense system is flexible. It is able to respond accurately to changes in the stimulus situation. The pathological ego, on the other hand, is much more rigid in its defense system. In the first place, because of unconscious needs, it is less sensitive to changes in its stimulus situation. Secondly, it is defective in integrative capacity because it has not been able to develop or maintain an adequate inter-relationship in its separate patterns.

Psychoneurotics may be conceived of as individuals who have more rigid defense systems than normals, they have some sensitivity to variations in situations which confront them and some capacity to modify responses in accordance with the differences in the stimulus situation. Therefore, we may expect a group of psychoneurotics taking a retest in which the conditions under which the test is administered are modified, to be able to shift more than a group taking a retest in which every attempt is made to have the conditions under which the retest is administered the same as the conditions under which the original test was given.

---

[1] We owe this formulation to M. L. Hutt. It was expressed by him in personal communication and in one of his publications. (12)

We may also expect that a group of psychoneurotics who improve in treatment will shift more in the experimental retest, than a group of psychoneurotics who do not improve in treatment. The psychoneurotic who is able to restructure his perception of a visual stimulus in accordance with a change in the test situation is expected to have a greater capacity for restructuring his perception of his "self" and the significant persons in his life and to be more responsive to therapy than neurotics not so endowed. The capacity to shift is conceived of as an indication of adaptability in life situations and as a differentiating criterion of mental health.

We conceive of the psychoneurotic as being less sensitive to change because of being unable to develop or maintain an adequate interrelationship in his separate patterns. This conception leads us to expect that a neurotic group taking a retest in which the conditions under which the test is administered are modified, will be less able to shift than a normal group given a retest under similarly modified conditions.

As a supplemental exploration, we have studied the shifts in other Rorschach variables which we correlated with the shift in the M factor. In this area of exploration, we have not tested any particular hypotheses. We investigated other variables in an attempt to learn more about the nature of the capacity to shift. It was expected that in general, the neurotic who is able to shift in the test

situation would be the person whose personality had departed less from the normal than the neurotic who was unable to shift. Therefore, it was expected that the shift group would show a greater number of the responses usually associated with healthy, normal records. It was expected that the shift group would have a greater number of M responses on the initial test. Similarly, the greater shift group was expected to show fewer of the pathological signs. It was expected that many of these differences would not be statistically significant. One cannot expect the predictive power of any single Rorschach sign to be very great. The absolute nature of a sign is almost meaningless for predictive purposes. A sign acquires significance only in relationship to other variables in the Rorschach psychogram. An attempt was made to develop a prognostic index for capacity to shift, based on the initial Rorschach pattern of those who shift more compared with those who shift less. This index was compared with the indices developed by Gibby (9), Milton (19), and Pottharst (22).

CHAPTER III

EXPERIMENTAL DESIGN

1. Subjects

All subjects were applicants for treatment at the Wayne County Mental Health Clinic, where this investigator is employed as one of the two staff psychologists. The clinic was opened in December, 1948, for treating adult ambulatory patients with personality disorders and as a measure for the prevention of more serious mental illnesses. A description of the structure of the clinic and an evaluation of certain aspects of therapeutic experiences at the clinic during its first year of operation is published (15). All subjects were selected only after meeting the following criteria. 1) Diagnosis by the intake psychiatrist as psychoneurotic not complicated by organic factors; 2) Must have at least two years of high school education; and 3) Must fall between the ages of 17 and 35 years. In order to avoid any bias in the selection of our sample population, we anticipated using all applicants meeting these three criteria until the desired number of subjects was obtained. This plan could only be approximated because a number of highly anxious persons made application for treatment and began treatment at a time when no psychologist was available for testing.

The following are the rationale for the selection criteria. 1) The healthy organism is conceived as being sufficiently sensitive to variations in situations which confront him, and sufficiently adaptive to these variations, to be able to modify his responses so as to deal effectively with them. The pathological ego is conceived as being much more rigid in its defense system. The psychoneurotic should fall somewhere between the healthy organism and the psychotic, in his ability to modify his responses to deal effectively with variations in situations which confront him. The non-compulsive psychoneurotic should be better able to alter his perceptions in accordance with altered directions than the compulsive psychoneurotic. It was felt that a group of psychoneurotics would be particularly well suited for a study of perceptual shifts in accordance with altered directions. Therefore, persons suffering from organic brain damage and those diagnosed as psychotic were excluded. The capacity to shift on the part of organics is another problem area. 2) The age restrictions are imposed to make the age of this group roughly comparable with the Milton group. It is desirable to limit this study to young adults as that is the group with which we are particularly concerned and we wanted to avoid completely involvement with problems of aging. 3) The minimum educational requirement was also employed to make the group roughly comparable to Milton's group. This criterion also served to exclude mental defectives or those of borderline intelligence. Intelligence or achievement test

scores were not available for most of our subjects.

Few applicants were excluded because of the age and education criteria. Almost all of the patients at the clinic were young adults with a high school education or more. Many potential subjects were excluded by the first criterion. A number of applicants were frankly psychotic or borderline psychotic or had a psychotic episode in the past. A number had a history which included some organic trauma or disease or convulsive disorders. Our criteria insured that we were dealing with a relatively homogeneous group of young adults who had at least ten years of formal schooling, who were diagnosed as suffering from a psychoneurosis, who had no present signs or known history of psychosis or organic brain damage and who were seeking psychiatric help at a public clinic for their difficulties.

2. Procedure

All subjects were given a Rorschach test according to standard directions, as soon as could be arranged after completion of the social worker's and psychiatrist's initial interviews.

The psychiatrist completed an initial rating form containing nine questions relating to severity of illness, prognosis, and compulsivity. Subjects were assigned to control and experimental groups on a random basis. Two weeks after the initial Rorschach, all subjects were re-examined by the original examiner. The 14 subjects

in the control group were again tested under standard directions. The 26 subjects in the experimental group were re-examined under instructions to report "everything you see and pay particular attention to people doing something." (Appendix I) Eight subjects in the control group were administered a specially created "testing the limits for M" (Appendix II) immediately after the retest. Twelve of the patients in the experimental group were given the M limits procedure immediately after the experimental set retest. [3]

Subjects were assigned to therapists (other than the investigator) by the psychiatrist-in-chief in the usual fashion as soon after the testing as could be arranged. When patients terminated treatment, the therapist completed a form for the evaluation of therapeutic change designed for this study. Exceptions to this procedure were those cases still in treatment at the time set for the analysis of the data. All of these subjects had six months or more of therapy. Treatment time of the terminated cases was sometimes under six months. Most subjects were seen

---

[3] In the original design, a third standard Rorschach was to have been given those subjects in both the control and experimental group who did not take the M limits test after six months or more of psychotherapy. This procedure was desired a) to provide a Rorschach measure of personality change and improvement in therapy, and b) to provide data on the stability of the various Rorschach determinants over a longer period of time. However, the original examiners were not available for this third Rorschach, and it was believed that using a second examiner would introduce an additional variable for which it would be difficult to control or correct. Had we known that the third Rorschach would not be given, all 40 subjects would have been given the M limits procedure.

twice weekly for 50 minutes of "psychoanalytically oriented" psychotherapy at each session.

3. Examiners

All testing was done by graduate psychologists other than the investigator. This was done to eliminate the conscious and unconscious bias of the investigator, had he done the testing. The examiners were not acquainted with the particular hypothesis under investigation. Sixteen subjects were tested by James M. Joyce, staff psychologist at the clinic. Nine were examined by Richard N. Filer as part of his doctoral field work. The remaining 16 subjects were tested by graduate students experienced in Rorschach test administration. The retest for each subject was given by the same person who gave the original test.

4. Scoring

Just as any possible bias of the investigator had to be eliminated in gathering the Rorschach protocols, conscious or unconscious bias also was eliminated in the scoring of the protocols. This was accomplished by having a person who was not involved in the research remove all identifying information from each record and assign it a number from a table of random numbers. This number, together with the necessary identifying data, was placed on a file card, which was not attached to the test protocol until all scoring was completed. Thus, there was no way for the scorer to know the experimental conditions under which the particular record

was obtained. Any errors in scoring can be assumed to be randomized. All scoring was done by the experimenter according to Beck's technique.

It is believed that the scoring attains the level of objectivity and reliability which a competent examiner may be expected to have. The scorer had extensive experience with the Rorschach and it can safely be assumed that his scoring accuracy is similar to that of most experienced Rorschach workers. Some of the variables were scored quite objectively, for example, those representing location. For other variables (F), there are tables which were used as a guide. Responses were scored as M only when they consisted of humans, including depersonalized humans, such as skeletons or cartoon figures, in some activity or animals in human-like activity. But for the purpose of this study, the most important point is that if errors in scoring occurred, they were <u>randomized</u>. There was no way for the scorer to know the experimental conditions for any record.

CHAPTER 4

ANALYSIS OF DATA

In the first three chapters we presented a statement of the problems to be investigated, the hypotheses to be tested and a description of the design of the experiment, including the detailed facts about the selection of subjects, the procedure, the examiners, and the scoring of the Rorschach protocols. Several additional studies are reported which were not anticipated in the original plan of the experiment, but which were carried out in an attempt to improve certain measures and to clarify the significance of some of the findings.

A total of 40 subjects was involved in this experiment. All 40 met the criteria given in Chapter 3 and had been accepted for treatment at the Wayne County Mental Health Clinic. All were given two Rorschachs approximately two weeks apart. The second test for the 14 subjects in the control group was given with standard directions, while the second test for the 26 subjects in the experimental group was given with altered directions designed to give a set for producing human movement responses. Of the 26 subjects in the experimental group, 12 were given an M limits test immediately following the retest Rorschach. Eight of the

14 in the control group were given the M limits test.

1. Initial Rorschach Factors in Control and Experimental Groups

In order to investigate the principal hypothesis of this study, it was first necessary to determine whether the control group and the experimental group were initially comparable groups on the experimental variable and on other relavent variables. Despite the efforts to assign subjects to the two groups on a random basis, this test of comparability had to be made. Any significant change between the two groups that would occur on the retest and that might be correlated with the size of the original score might not then be attributable to the experimental conditions but to discrepancies between groups on original test scores. The statistical data and results of these comparisons for the experimental variable, human movement, and for the weighted sum of the color responses, and for the total number of responses, are shown in Table 1. The data for sum C is presented because M is generally interpreted clinically not by itself but in relation to other variables, and in particular to the sum C in the M : Sum C ratio, which is known as the experience balance from the German "erlebnistypus". The abbreviation "E.B." for "experience balance" will be used from this point. The data for the number of responses is included because this factor must be considered in the evaluation of any Rorschach variable. The same number of M responses in Rorschach records of widely differing length,

may have clinically different meanings. If the two groups had differed significantly in R, it would have been desirable to evaluate M change in terms of per cent of M change as well as in terms of change in number of M.

Table I shows that for mean number of human movement responses and for mean number of responses, the t's did not attain statistical significance at less than the 10% level of confidence. The mean weighted sum of color responses did attain statistical significance at less than the 10% level. Therefore, we conclude that the control group and the experimental group can not be regarded as comparable in mean Sum C. The control and experimental groups are comparable in mean initial M and R. It should be noticed that the mean M's for our two groups are 2.00 and 3.42. Rorschach workers generally report two to four M's as average for normal subjects and they report fewer M responses for neurotics. It is suggested that our sample may not be typical for all neurotics and therefore our findings may have limited applicability.

2. M Set Change in Control and Experimental Groups

After finding that the control and experimental groups were comparable with respect to mean M and mean R in the initial test, the second task was to determine whether or not a neurotic group was able significantly to increase its perception of human movement in a retest under an externally given set to see human movement. Inspection of the

TABLE 1

COMPARISON OF INITIAL TEST OF CONTROL GROUP WITH INITIAL TEST OF EXPERIMENTAL GROUP FOR SELECTED RORSCHACH VARIABLES [1]

Variable	Control N=14 Mean	S.D.	Experimental N=26 Mean	S.D.	Diff. Between Means	S.E. of Mean	t	P
# of M	2.00	1.71	3.42	3.97	1.42	1.12	1.27	.20-.30
Sum C	2.07	1.47	4.42	3.47	2.35	.96	2.45	.02
R	30.8	17.8	32.0	20.8	1.2	1.53	.78	.40-.50

[1] see Appendix IV, page 110.

distributions of M scores on the initial test reveals that both distributions were skewed. The distribution resembles one half of a normal distribution. It is suggested that this curve is produced in part because of the absence of negative scores. Skewness is produced by high frequencies of zero scores. The subjects with zero scores may be psychologically different. These scores are alike, because M is a crude measure. An analysis of animal movement and inanimate movement in their Rorschach protocols might be fruitful, but that is beyond the scope of this study. Even though the control and experimental groups are statistically comparable in mean initial M nevertheless there are differences in the initial M scores for the two groups which may be clinically significant.

It was decided that the most rigorous and accurate test of the significance of the change in M in the experimental group and in the control group would be to determine whether or not the mean of the differences between M on the initial and repeat test for each group differs significantly from zero.

The statistical data for answering this question are presented in Table 2. It shows that there is no significant difference in M between the initial and repeat test for the control group and that there is a significant difference in M beyond the 1% level of confidence between the initial and repeat test for the experimental group. It will be recalled that the two groups are comparable with respect to initial

mean M. It is evident that M remains constant in a standard test, retest situation after a short interval. The change in the experimental group is not a result simply of a repeat test. We infer that the shift in M in the experimental group is due to the externally imposed set under which the repeat test was given to the members of the experimental group. We may safely conclude that neurotics are able to increase their perception of human movement under the conditions of this experiment.

The average increase in M in the experimental group from 3.4 to 7.4 is not only statistically significant, but is clinically significant as well. For normals, at least, approximately 2 to 4 human movement responses per record is generally considered to be about average, while 7 or 8 M's would definitely be regarded as high M productivity. A Rorschach protocol containing 3 or 4 M responses would yield a different personality interpretation than a record with 7 or 8 M responses. Our findings indicate that for our experimental group, the perception of human movement is not necessarily a "deep" unconsciously determined variable, but one which can easily be altered under changes in the external situation. This finding is contrary to Rorschach and authorities since Rorschach.

However, M is generally interpreted in relation to the responsiveness to color, as well as by itself. The changes in this M : Sum C ratio is presented in the next section.

TABLE 2

COMPARISON OF HUMAN MOVEMENT RESPONSE ON INITIAL TEST WITH REPEAT TEST FOR CONTROL AND EXPERIMENTAL GROUPS [1]

Group	N	Mean M Initial	Mean M Repeat	Diff. Between Means	t	P
Control	14	2.00	1.79	-.21	.80	.50 - .60
Exper.	26	3.42	7.42	4.00	4.33	.01

[1] See Appendix 4, page 110.

TABLE 3

CHANGES IN EXPERIENCE TYPE FOR CONTROL AND EXPERIMENTAL GROUPS

Group	Experience Type	ft* Initial	fo Repeat	X^2	P
Control N=14	Introversive	3	4		
	Extratensive	3	4		
	Ambi. and Co-art.	8	6	1.166	.50-.60
Experi. N=26	Intraversive	7	16		
	Extratensive	13	3		
	Ambi. and Co-art.	6	7	19.415	.01

* The theoretical frequencies were taken from each group separately rather than from both groups combined, which is customary, because it was believed that the former method provided a more vigorous test of the hypotheses.

3. Change in M : C Ratio of Control and Experimental Groups

In order to study the changes in the M : C ratio, subjects were placed in four experience balance groups. For this purpose, responses were weighted in the standard fashion: each human movement response was weighted one M, while the color responses were weighted ½C for each response scored FC, 1C for each response scored CF, and 1½ C for each response score C. The four E.B. groups were defined, in accordance with prevalent practice, as follows:

1. Introversive: M greater than Sum C plus one.
2. Extratensive: Sum C greater than M plus one.
3. Co-artative: M and Sum C both one or less.
4. Ambiequal: Either M or Sum C greater than one and the difference between them one or less.

Two hypotheses were tested: a) the control repeat test produces no statistically significant change in the experience type in comparison with the control initial test; and b) the experimental repeat test produces no statistically significant change in the experience type in comparison with the experimental initial test. The results of the statistical tests of these hypotheses are shown in Table 3. In this table the theoretical frequency (Ft) is the number of cases in each experience type in the initial test, and the observed frequency (Fo) is the number of cases in each experience type in the repeat test. The co-artative and ambiequal groups are combined in order to make the theoretical frequencies as large as possible without violating the

essential test which is the change in the introversive-extroversive balance.

Table 3 shows that the control group does not change its M : C ratio. The experimental group does change its M : C ratio in the direction of introversiveness. This change is significant beyond the 1% level of confidence. We conclude that in this study the E.B. does not change on a simple repeat test and that the E.B. does change in the experimental repeat test.

An examination of the records of the individual subjects shows that of the 14 cases in the control group, 9 show no change in E.B. type. The other 5 subjects showed minor changes in E.B. type. None shifted from introversive to extratensive or from extratensive to introversive.[1]

The experimental group shows very different results. Of the 26 subjects in this group, 12 show no change in E.B. type. Five change from extratensive to introversive and five others from extratensive to either ambiequal or co-artative. In other words, all subjects in the experimental group whose E.B. was changed, changed in an introversive direction. They changed in the direction of higher M in the M : C ratio.

---

[1] one from ambiequal to introversive
one from ambiequal to co-artative
one from introversive to ambiequal
one from extroversive to co-artative
one from extroversive to ambiequal

4. M : C Ratio in Normals and Neurotics

Now we will compare the changes in M and in E.B. in the neurotic groups of this study with the normal groups of Milton's study. It is difficult to compare the changes in M and in E.B. directly because of the marked differences in the scores in the variables in the initial test. The mean initial M, mean repeat M, mean gain in M, and mean gain in M as a percentage of mean initial M for the normal experimental group and the neurotic experimental groups are:

Group	Initial M	Repeat M	Gain	% Gain
Normal Experimental	6.85	13.61	6.76	99
Neurotic Experimental	3.42	7.42	4.00	118

It is evident that the normal group is markedly higher in mean initial M and shows a higher mean gain in terms of raw M score. It is also evident that the neurotic group, in terms of percentage gain shifts fully as much as the normals. The conclusion is suggested that both the normal and neurotic groups were able to increase their mean M score with comparative ease. It may be of some interest to note that the range of change in M for the normal experimental group is -1 to +17, while the range of change in M for the neurotic experimental group is -3 to +17. The range of shift between the two groups is much more comparable than the mean initial M and the mean gain in M.

More definite and significant conclusions about the comparability of the capacity to shift in the normals

compared with that in the neurotics were obtained by studying the shift in the experience balance. The E.B. is given more weight in interpreting a Rorschach record than is given the M score alone.

In order to compare the shift change in E.B. between the normals and neurotics it was first necessary to analyze our data in the way Milton analyzed his data. We categorized our subjects in terms of M greater than C, C greater than M and M and C equal, as Milton did. The data for Milton's groups and for our groups are presented in Table 4.

TABLE 4

THE NUMBER OF SUBJECTS IN EACH GROUP FOR BOTH ORIGINAL AND REPEAT TEST WITH THE SPECIFIED DIRECTION OF THE E.B.
(A Comparison of Milton's Data* with the Data of the Present Study.)

Group	Original			Repeat		
	M greater	C greater	equal	M greater	C greater	equal
Normal Con.*	10	11	0	13	7	1
Neurotic Con.	6	7	1	6	6	2
Normal Exp.*	9	11	1	19	1	1
Neurotic Exp.	8	14	4	21	5	0

To analyze the data in Table 4 several chi square tests were made. Milton obtained a chi square of 4.59, significant between the 2 and 5 per cent level, for the relationship between membership in his control and experimental groups and shift in any direction in the E.B. We made a similar test for our groups. (Table 5)

TABLE 5

RELATIONSHIP BETWEEN MEMBERSHIP IN A GROUP AND SHIFT IN DIRECTION OF E. B.

	Shift	No Shift	
Neurotic Control	2	12	14
Neurotic Experimental	14	24	26
	16	24	

Chi square = 5.94

P = .01-.02

It will be noted that within both the normal and neurotic groups there is a significant difference in shift in direction of the E. B. for the experimental compared with the control groups.

Two additional statistical tests revealed that there is no statistically significant difference between the normal and neurotic control groups or the normal and neurotic experimental groups in number of subjects shifting in direction of E. B. (Table 6 and 7.)

TABLE 6

RELATIONSHIP BETWEEN MEMBERSHIP IN NORMAL CONTROL AND NEUROTIC CONTROL GROUP AND SHIFT IN DIRECTION OF E.B.

	Shift	No Shift	
Normal Control	6	15	21
Neurotic Control	2	12	14
	8	27	

Chi square = 1.312

P = .20-.30

TABLE 7

RELATIONSHIP BETWEEN MEMBERSHIP IN NORMAL EXPERIMENTAL AND NEUROTIC EXPERIMENTAL GROUP AND SHIFT IN DIRECTION OF E.B.

	Shift	No Shift	
Normal Exper.	13	8	21
Neurotic Exper.	14	12	26
	27	20	

Chi square = 0.285

P = .50-.70

These findings suggest that the neurotics are able to shift the direction of the E.B. as readily as the normals. The capacity to shift in the direction of the E.B. does not differentiate the two groups. This finding does not support the concept of capacity to shift on the Rorschach as a criterion of mental health. If the normals are a less pathological group than the neurotics, a reasonable assumption to make on the basis of the manner in which the two groups were selected, then the normals should shift more readily. It is suggested that capacity to shift as here measured may not be an indication of a healthy type of flexibility. Possible meanings for this finding are discussed in the next chapter. However, the fact that our experimental neurotic group had a mean initial M of 3.42 suggests that it is probably not comparable to all neurotic groups. Comparable results might not be obtained if this experiment were repeated with another neurotic group.

5. Relationship Between Initial M and M Set Change

The next task is to see what the relationship is between the number of human movement responses on the initial test and M set change. If initial M and M set change are perfectly correlated then no increase in the prediction of therapeutic change from M could be obtained from the M change data. When the 26 patients in the experimental groups are dichotomized on initial M into the 13 highest on M and the 13 lowest on M and are compared on M change, X^2 proves to be 1.418, which is not significant. This value for X^2 falls between the 30 and 50% level of significance. In other words, there is no significant difference in M change between patients low on initial M compared with patients high on initial M.

An examination of the scattergram of initial M and M set change suggested the existence of a curvilinear relaship. The scattergram indicated that subjects low in initial M tend to give a moderate increase in M in the retest; individuals high in initial M tend to fall into two distinct sub-groups, a) those who give very high increase in the retest, and b) those who give zero and negative M change in the retest. The range in M change was found to be from minus 3 to plus 17, with a mean change of plus 4.0. For testing this suggested curvilinear relationship, M change was divided into 3 groups (low = minus 3 to 0, with 7 patients; moderate = plus 1 to plus 7, with 13 patients;

high = plus 8 to plus 17, with 6 patients). Initial M was divided into two groups (low = 0 to 2, with 15 patients; high = 3 to 16, with 11 patients).[1] The X^2 for this relationship proved to be 15.197, which is significant beyond the 1% level of confidence.

The small sample limits the decisiveness with which conclusions can be made. Our data supports the conclusion that, within the experimental groups, the relationship between initial M and M set change is not a linear one. Subjects who are low in initial M (0 to 2) tend to give a moderate increase in the repeat test (change of +1 to +7). Subjects who are high in initial M (3 to 16) fall into two groups: a) those who give very high increase in the repeat test (change of plus 8 to plus 17), and b) those who give no increase or decrease in the repeat test (change of minus 3 to zero).

An exploratory study of the differences between the high initial M patients who increased M in the retest compared with those who did not increase M is reported in section 7 of this chapter.

6. Stability Scores

In this section we want to determine whether certain Rorschach characteristics in the initial test can distinguish between those subjects who shift most in M from those who shift least. Gibby (9), Milton (19), and Pottharst(22),

[1] This division was made on the basis of the distribution of initial M scores and M change scores.

each developed formulae, called stability scores, to predict the direction and amount of change in a particular Rorschach factor. Pottharst presents a comparison of the three stability scores. He says, "two out of three stability scores (Milton's and Pottharst's) seem effective only under conditions of the specific set given the subjects, and the third score (Gibby's) seems effective in predicting under conditions of two simple sets and ineffective in predicting in a complex set. This points to the conclusion that the specific set held by the subject is of crucial importance in determining the type of change in Rorschach scores." Pottharst observes further, "also, it should be remarked that the scores developed by Milton and Gibby have more in common with each other than they have with the one developed (by Pottharst)."

We applied Milton's score to our data because it is the only score developed for subjects who were given the same set as our experimental subjects and it seems to have some generality of application with respect to normal subjects. Milton found that his shift most group showed significantly more of the sum of the following characteristics on the initial Rorschach, than his shift least group: 1) W minus M score of plus one to plus seven; 2) CF score of three to five; 3) S score of three to five; 4) Sum Y score of below six; and 5) W minus Sum C score of zero or minus value.

We dichotomized our experimental group into two subgroups, 13 subjects who shifted most (+3 to +17) and 13

subjects who shifted least (-3 to +2), on the repeat test. The following chi square and P values were obtained. The group showing the greater frequency for each item is indicated. (Table 8).

Examination of Table 8 reveals that three of the five factora differentiate between our most and least shift group beyond the 10 per cent level of confidence. However, two of these three factors (W minus M of plus one to plus seven and Sum Y below six) characterize our <u>least</u> shift group, while they characterize Milton's <u>most</u> shift group. Only one factor (S of 3 to 5) characterizes both the normal and neurotic shift most group. The evidence suggests that Milton's stability score does not apply to our group of neurotics.

The next step was to develop a stability score for the experimental group in the present study. We chose to use the same initial Rorschach items as Milton used. Examination of the distribution of scores in these five items showed that, in general, the shift most sub-group was more variable than the shift least sub-group. Therefore, we developed a stability score of predicting (really post-dicting) which subjects would shift least, or, in other words, be stable in productivity of M between the initial and repeat experimental test. This new stability score included four of the five Rorschach factors used by Milton, but with different numerical values for each of them. The new

stability score factors with chi square and P values for our experimental subjects dichotomized in the same way as for the previous test of Milton's stability score, is given in Table 9.

Examination of Table 9 shows that each of these four factors effectively discriminates between the least and most shift group beyond the 10 per cent level of confidence; three of them beyond the 5 per cent level and two of them beyond the 2 per cent level of confidence.

Stability scores were then computed for all 26 experimental subjects using the new stability score factors and the most shift and least shift groups were compared. Table 10 shows the data for the two sub-groups.

Inspection of Table 10 shows that the t for the difference in stability score between the most shift and least shift group is statistically significant beyond the 1 per cent level of confidence.

As these results were obtained on only one group of 26 subjects, we are not able to state whether or not they would be applicable to other groups. However, the hypothesis that there is no difference other than that attributable to chance, in our stability score of those individuals in this experiment who changed the most compared with those individuals who changed the least in M under the M set conditions, can be confidently rejected. The generality of this finding would depend on the outcome of a rigorous cross-validation study in which predictions are made as to

TABLE 8

MILTON'S STABILITY SCORE FACTORS WITH CHI SQUARE AND P VALUES FOR NEUROTIC EXPERIMENTAL GROUP DICHOTOMIZED IN MOST SHIFT (N=13) AND LEAST SHIFT (N=13) SUB GROUPS

Factor	Chi Square	P	Higher Score Sub Group
W - M +1 to +7	2.889	.10	least shift
CF 3 to 5	0.258	.50 - .70	most shift
S 3 to 5	2.728	.10	most shift
Sum Y below 6	4.960	.05	least shift
W - Sum C 0 or -	0.432	.50 - .70	most shift

TABLE 9

GOERTZELS STABILITY SCORE FACTORS WITH CHI SQUARE AND P VALUES FOR NEUROTIC EXPERIMENTAL GROUP DICHOTOMIZED INTO MOST SHIFT AND LEAST SHIFT SUB GROUPS

Factor	Chi Square	P
W - M +1 to +6	2.888	.10
S 0 to 2	6.190	.02
Sum Y less than 2	6.050	.02
W - Sum C 0 to +2.5	4.260	.05

TABLE 10

COMPARISON OF STABILITY SCORES OF SUB GROUPS OF EXPERIMENTAL GROUP DICHOTOMIZED AS TO MOST AND LEAST CHANGE IN M SET RETEST

Subgroup	Mean	S.D.	Mean Diff.	S.E. of Mean Diff.	t
Most shift	1.54	.93	1.46	.15	9.07*
Least shift	3.00	.96			

*t for the 1 per cent level is 2.80

direction and amount of shift on the basis of the stability score on the initial Rorschach.

It is difficult to interpret these findings in psychological terms because the psychological significance of various frequencies of the different Rorschach determinants and relationships between determinants has not been conclusively demonstrated. If Milton's concept is correct, namely, that W minus M scores of plus 1 to plus 7 reflect greater ego strength than other values of W minus M, then, in our group, the patients with greater ego strength shift less than patients with lesser ego strength. One could theorize that the stronger patients were better able to resist the set and keep to their original responses with respect to M. The psychological meaning of the S and Sum Y factors, both of which are low for the least shift group, is difficult to determine. The difference in total number of responses must be taken into consideration, even though these determinants are usually evaluated from the raw score rather than from the percentage of total responses. For our two sub-groups there is a marked difference in average number of responses on the initial Rorschach. In fact, R can be used as another factor in a stability score. The chi square test for R below 30 compared with R of 30 and above, distinguishes between the shift most and shift least group. (Chi square = 3.74, P = .10). The mean R for the most shift group is 42.4, compared with a mean R of 21.5 for the least shift group.

From one viewpoint, the difference in R for the two groups may suggest that the richer personality shows more shift than the more impoverished personality. However, in view of the other findings, particularly the data for W minus M, it may be that the large number of responses in the shift most group is a reflection of a compulsive trend in the shift most subjects. The large R may stem from a drive toward thoroughness or may be an attempt to compensate quantitatively for whatever limitations the individual may feel within himself.

Looking at the R factor in this latter manner, the results would be congruent with the finding and interpretation offered for the W minus M factor. We might theorize that both findings point toward the acceptance of the set offered by the examiner in the experimental repeat test as behavior shown by individuals with weak ego strength who use compulsive defenses. The low number of Y responses may reflect less anxiety in the shift least group than in the shift most group. The Y findings are also congruent with this speculation.

However, S is generally regarded as reflecting oppositional tendencies. The least shift group show little or no S. On the face of it, it would seem that the patients who shift least are less oppositional than patients who shift most. But, if they are less oppositional, why don't they accept the set given the examiner? Perhaps, it is because they are unable to see more human movement in the ink

blots. No final conclusions can be offered by the writer. It is suggested, however, that the same patterns on the initial Rorschach have different psychological significance, in terms of M change between a standard Rorschach and an M set retest Rorschach, for neurotics than for normals. For the normals in the Gibby, Milton, Pottharst experiment, all three stability scores, in a general sense, suggested that richness, emotional responsiveness and lack of constriction were characteristics of normal people who can shift in Rorschach score. These same personality characteristics do not appear to be present in neurotics who can shift in Rorschach score.

7. Exploratory Study of Differences Between High Initial M Patients Who Increased M In Repeat Test and Those Who Did Not Increase M In Repeat Test

This exploratory study was undertaken to see what differences might exist between the initial high M patients who increased M in the repeat test and those who did not increase M in the repeat test. It was suggested that the highly significant statistical difference in capacity for increasing the perception of human movement in the experimental repeat test might be related to personality characteristics which would be reflected in the initial Rorschach variables and might be related to the capacity to change under psychotherapy.

First it was determined that the difference in M set change is not related to change in therapy as rated by the therapists. Of the seven patients in the high increase group, four improved and three did not improve in treatment. Of the five patients in the low increase group, three improved and two did not improve in treatment.

Inspection of Table XI, shows that the difference in change in M between the two groups is significant beyond the 1% level of confidence. The high increase group showed a mean shift in M of plus 10.0 while the low increase group showed a mean increase of 0.2. It is clear that the two groups are very different in their amount of shift. It was

on this basis that the two groups were selected. The high increase group gives a larger number of total responses on the initial Rorschach than the low increase group. This difference is significant at the 5% level of confidence. It also will be noted that the high increase group is higher in the initial test in mean M, mean Sum C, and mean Sum SH. (Sum SH is obtained by weighing the shading responses, which include Y and V responses, in the same way as color responses are weighed to get Sum C). However, none of these differences is significant at the 10% level of confidence, despite the large size of the differences between the means of the two groups. This lack of statistical significance may be attributed to the small size of the sample and the wide dispersion of scores within each group. Nevertheless, the data suggested that there is a difference between the two groups in terms of constriction, responsiveness, and richness or complexity of the personality structure.

A "richness" score was obtained to see whether or not our concept, that the high increase group has a richer personality structure as reflected in the initial Rorschach, would be supported statistically. The richness score is the sum of M, plus Sum C, plus Sum SH. In other words, the richness score represents the total of all the determinants (according to Beck's system of scoring) used in making the responses to the Rorschach figures, except form. Although the high increase group has a mean richness score of 19.7 compared to 8.6 for the low group, the difference is not

quite significant at the 10% level of confidence. (t at 10% level is 1.812; obtained t is 1.78).

There is a trend in our findings which suggests that high M patients who are able to increase M in an M set repeat test are richer, more complicated personalities than high M patients who are unable to increase M in the repeat test. Our findings suggest that the ability to increase M in high M patients is not related to improvement in psychotherapy.

We are not justified in drawing any conclusions because of the lack of decisiveness of the results and the small sample in this exploratory study which was not included in the original plan of the experiment. A study specifically designed to study the relationship of the ability to increase the perception of a Rorschach determinant in a group high in that variable in the initial test, to various aspects of the personality and to change in therapy, is indicated.

8. Psychiatrists Initial Rating and M Set Change

An initial rating form containing nine items was adapted from a scale devised by Hutt[1] and found usable by Schrier (24). Our form was designed to facilitate the quantification of the psychiatrists judgments of severity of illness, of prognosis and of compulsivity which they made at the time of their initial interviews with the patient. The

---

[1] This form was developed by M. L. Hutt in connection with his research on psychotherapy.

ratings were completed before treatment began. A comparison of the scores given to patients[1] by the different intake psychiatrists suggests that they were using similar standards and frames of reference in making their evaluations. The means and the ranges of scores of the patients judged by the three psychiatrists who rated 36 of the 40 patients in this study, are almost identical.[2]

The initial rating form was completed by the intake psychiatrist in an attempt to evaluate the relationship between M change and severity of illness, prognosis, and compulsivity as judged by the psychiatrist, before treatment began. It was expected that patients rated as more compulsive would be less able to shift on the retest Rorschach than patients rated as less compulsive. In order to make statistical evaluations, numerical values were assigned to the ratings on an ad hoc basis, as noted in Figure I.

---

[1] Originally the intake psychiatric social worker also judged the patients on the initial rating form. The social worker and the psychiatrist did not agree very well in their judgment of the patients. Therefore, only psychiatrists' ratings were used in this study, as they were assumed to have more face validity. However, the low reliability of the scale, inferred from the relatively low positive relationship between social worker and psychiatrist ratings of the same patients (based on each staff member's own interview), tends to attenuate our measures of relationship between the patient characteristics assessed from the initial rating form and other variables.

[2]
Dr. 1 mean = 23.0 Range 18 - 29
Dr. 2 mean = 22.7 Range 17 - 28
Dr. 3 mean = 22.5 Range 19 - 28

TABLE 11

COMPARISON OF TWO GROUPS OF HIGH INITIAL M SUBJECTS:*
H HIGH INCREASE M ON REPEAT TEST (N = 7)
L LOW INCREASE M ON REPEAT TEST (N = 5)

Variable	Group	Mean	Diff. Between Means	t	P
M change	H	10.00			
	L	0.20	9.80	5.44	.01
Initial R.	H	45.71			
	L	24.20	21.51	2.32	.02 - .05
Initial Richness Score	H	19.7			
	L	8.6	11.1	1.78	.10 - .20
Initial Sum SH.	H	6.79			
	L	2.20	4.59	1.77	.10 - .20
Initial Sum C	H	6.50			
	L	2.60	3.90	1.64	.10 - .20
Initial S	H	2.43			
	L	0.80	1.63	1.44	.10 - .20
Initial F-	H	7.71			
	L	4.20	3.51	1.23	.20 - .30
Initial M	H	6.43			
		3.80	2.63	1.05	.30 - .40
Initial P	H	6.86			
	L	4.80	2.06	.94	.30 - .40
Initial M-	H	1.71			
	L	0.80	0.91	.70	.50

* The reader should be aware that the scores dealt with in this table are not independent scores.

FIGURE I

INITIAL RATING FORM USED BY INTAKE PSYCHIATRISTS WITH WEIGHTS GIVEN TO EACH ITEM

	slight	moderate	marked	extreme
1. Extent of underlying pathology. (e.g. severity of neurosis.)	1	2	3	4
2. Extent to which symptomatic manifestations of neurotic illness is:				
a) personally incapacitating	1	2	3	4
b) socially incapacitating	1	2	3	4
3. Likelihood that patient will be able to modify the severity of the underlying pathology.	4	3	2	1
4. Likelihood that patient will be able to modify the severity of the symptomatic manifestations of neurotic illness.	4	3	2	1
5. Extent of rigidity of thinking processes. (e.g. difficulty in considering all aspects of a rpoblem because of conflict areas.)	1	2	3	4
6. Extent to which repetitiveness of behavior based on neurotic needs is:				
a) socially incapacitating	1	2	3	4
b) personally incapacitating.	1	2	3	4
7. Degree of compulsivity shown in mannerisms and speech.	1	2	3	4

Comments.

Statistical analysis revealed no significant relationship between the psychiatrists' overall initial rating (total score on initial rating form) and M set change. (see table 12)

TABLE 12

RELATIONSHIP BETWEEN SCORE FOR PSYCHIATRISTS' OVER ALL INITIAL RATING AND M SET CHANGE

(N = 26)

M Set Change	Overall Initial Rating			
	Low	Mod.	High	
High	4	2	2	8
Mod.	2	3	4	9
Low	2	4	3	9
	8	9	9	

Chi-square = 2.49
R = 50% - 70%

Similarly, the chi-square statistic revealed no significant relationship between the rating of severity of initial illness (total score for items 1, 2a and 2b) and M set change (see Table 13).

TABLE 13

RELATIONSHIP BETWEEN SCORE FOR PSYCHIATRISTS' RATINGS OF SEVERITY OF INITIAL ILLNESS AND M SET CHANGE

(N = 26)

M set Change	Initial Illness: Mod.	Sev.	
High	9	4	13
Low	5	8	13
	14	12	

chi-square = 2.49
P = 10%-20%

Although the relationship is not statistically significant and the null hypothesis of no relationship between degree of initial illness and M set change can not be rejected, there is a tendency in our sample for patients who are judged moderately ill to increase M more in the experimental repeat test than patients who are judged severly ill.

A similar chi-square test showed no statistical difference at all in M set change for patients judged slight and moderate in "extent of underlying pathology e.g., severity of neurosis" (item 1 of Initial Rating Form), compared with patients judged marked or extreme in extent of underlying pathology.

However, as expected, the rating of compulsivity given by the psychiatrists (total score for items 5, 6a, 6b and 7 or the initial rating form) is related to M set change. (See Table 14).

TABLE 14

RELATIONSHIP BETWEEN PSYCHIATRISTS' RATING OF COMPULSIVITY AND M SET CHANGE

(N = 26)

	Compulsivity			
M set Change	Low	Mod.	High	
High	4	1	1	6
Mod.	5	3	5	13
Low	0	5	2	7
	9	9	8	

chi-square = 8.81
P = .05 - .10

It is concluded that patients judged more compulsive tend to increase M less on the experimental repeat test than patients judged as less compulsive. In other words, in our sample and with the measures used, there is a tendency for a negative relationship between compulsivity and shift in M in the experimental situation.

A Pearson r was computed between "neurotic signs" on the initial Rorschach and psychiatrists' judgments of initial illness (total score for items 1, 2a and 2b on the initial rating form) to determine the relationship between these two methods of evaluating severity of neurosis. It was believed that a significant positive correlation between the two measures would lend confidence to the validity of both measures of severity of neurosis. The Rorschach neurotic signs used were six of the eight factors used by Miale and

Harrower-Erickson (discussed later in this paper on page 73ff.) These signs effectively distinguished between neurotics and normals in their study (17). We obtained an r of +.14, which is not statistically significant. This finding suggests that either or both of these measures might be invalid. The correlation of +.14 indicates that they are not related in this sample.

9. Evalutaion of Therapeutic Change

The problem of how to get a valid criterion measure of therapeutic change is a thorny one, which has not been solved. In studies of hospitalized patients, frequently the criterion is simply whether or not the patient is still in the hospital at a certain date after the experimental procedure (shock, psychotherapy, etc.) was completed. In studies of out-patients, generally the criterion is an overall evaluation by the therapist on some simple scheme, such as "definitely improved," "somewhat improved" or "unimproved." We felt that we could improve on such a gross system by the use of a more refined rating form which required the therapist to make a number of more specific judgments of pertinant aspects of therapeutic change. The use of objective criteria of improvement was rejected for this study because the case records at the clinic were inadequate for obtaining evidence of objective changes in the lives of the patients.

The rating form used in this study is a revision by this author, in cooperation with the staff at the Wayne County Mental Health Clinic, of a form previously developed by Hutt

and his associates at the University of Michigan. The clinic staff members were given some training in the use of the rating form in an effort to increase the comparability of their judgments. As a part of this training, one patient (not a subject in our experiment) whose case had been presented to the staff at therapy conferences for over one year, was rated by eight staff members, including the therapist. On the first ten items of the rating form (the portion used in this experiment) 78 per cent agreement was found among all eight raters (omitting the 10 don't know responses out of the total of 80 responses). On seven of the ten items, the majority of the judgments of the other raters was in agreement with the judgment of the therapist.

The form for the evaluation of therapeutic change which was used in this study, together with the numerical weights assigned to each item on a common sense basis for statistical analysis, is found in Figure 2 . The score for the therapeutic change for each patient is the sum of the weights, as given in Figure 2 , for each item checked by the therapist as applying to the patient. Items 11 and 12 were not included in the score for therapeutic change because it was believed that they would contaminate the scores. All the other items represent an aspect of the therapist's own evaluation of change in the patient. Item 11 refers to the patient's feelings and Item 12 refers to feelings of others than the patient and the therapist.

FIGURE 2

RATING OF THERAPEUTIC CHANGE FORM USED BY THERAPISTS WITH WEIGHTS GIVEN TO EACH ITEM

		Yes	No	Don't Know
1.	Has the ego been healthfully strengthened?	1		
2.	Is more energy available for the environment?	1		
3.	Has there been a reduction in symptoms?	1		
4.	Does the patient still us the same defenses?		1	
5.	Has the patient's insight increased?	1		
6.	Has there been improvement in sexual adjustment?	1		
7.	Has there been improvement in vocational or educational adjustment?	1		

		None	Slight	Mod.[1]	Marked	Don't Know
8.	Extent of change in underlying pathology. (e.g. severity of neurosis)		1	2	3	
9.	Extent of symptomatic improvement.		1	2	3	
10.	How improved is the adequacy of his (emotional) understanding of his total environment?		1	2	3	
11.	How much improvement does the patient himself feel he has made?*					
12.	How much improvement do others than the patient and the therapist feel the patient has made?*					

1 Moderate

* Items 11 and 12 were not included in the score for therapeutic change. It was believed that would contaminate

the scores. All the other items refer to aspects of the therapist's own evaluation. Items 11 and 12 require a judgment from the therapist as to the feelings of other persons.

FIGURE 3

Frequency of Response to Each Item on the Therapists' Rating of Therapeutic Change Form

		N_{Yes}	N_{No}	$N_{Don't\ Know}$
1.	Has the ego been healthfully strengthened?	8	17	8
2.	Is more energy available for the environment?	13	17	3
3.	Has there been a reduction in symptoms?	18	14	1
4.	Does the patient still use the same defenses?	21	9	3
5.	Has the patient's insight increased increased?	16	14	8
6.	Has there been improvement in sexual adjustment?	6	19	8
7.	Has there been improvement in vocational or educational adjustment?	9	14	10

		None	Slight	Mod.	Marked	Don't Know
8.	Extent of change in underlying pathology. (e.g.severity of neurosis.	20	10	1	0	2
9.	Extent of symptomatic improvement.	12	12	5	3	1
10.	How improved is the adequacy of his (emotional) understanding of his total environment?	15	12	3	2	1

Figure 3 (continued)

	None	Slight	Mod.	Marked	Don't Know
11. How much improvement does the patient himself feel he has made?	9	10	4	4	6
12. How much improvement do others than the patient and the therapist feel the patient has made?	11	6	2	2	12

A tabulation of the frequency of the response to each item for the total group of 33 subjects who were in treatment, and therefore could be rated, is given in Figure 3.

Examination of the ratings of therapeutic change for the experimental group showed great differences in the average score given to different therapist's patients. It was believed that certain aspects of the therapists' needs were influencing their ratings of the progress made by their patients. It was believed this criterion of therapeutic change was quite fallible, despite our efforts to develop a valid and reliable criterion. Nevertheless, statistical tests based on these ratings were made. Therapists' ratings were the criteria of therapeutic change selected in the design of the experiment inasmuch as therapists' ratings are the standard criteria for judging improvement in psychotherapy.

The therapists' ratings of therapeutic change was first used to see if there was any relationship between number of M responses on the initial Rorschach and Therapeutic

change. It was found that in our sample of 20 experimental patients who entered treatment there was no statistically significant difference in ratings of therapeutic change for patients with a low number of initial M responses compared to patients with a high number of initial M responses. When the distribution of number of M responses was dichotomized and the distribution of therapeutic change scores based on therapists' ratings, was dichotomized, the X^2 value was .833, which is not significant. This value for X^2 falls between the 30 and 50% level of significance.

Then we tested the hypothesis that patients with a moderate number of initial M responses would respond more favorably to treatment than patients who give no M responses or a great many M responses. (This is the concept of the optimum).

For testing this hypothesis, the number of initial M responses was divided into two groups (<u>moderate</u> = 1 to 3, with 9 patients; <u>extreme</u> = 4 and above or 0, with 11 patients). Therapeutic improvement was dischotmized into improved and unimproved with 10 patients in each group. The X^2 for this relationship proved to be 1.82, which is not significant. This value for X^2 falls between the 10 and 20% level of significance.

We concluded that within the limits of this study there is no simple statistical relationship between number of M responses on the initial Rorschach and therapeutic improvement as rated by the therapists. Patients high in initial M

do not improve in treatment more than patients low in initial M. Patients with a presumably optimum number of initial M responses do not show more improvement in treatment than patients with no M and with a very high number of M responses.

The basic hypothesis that there would be a positive relationship between M set change and change in therapy as reported by the therapist was not supported. When the distribution of M change was dichotomized and therapeutic change as rated by the therapists was dichotomized, for the total of 20 patients in the experimental group who entered treatment, the X^2 value was .800, which is not significant. This value for X^2 falls between the 30 and 50% level of significance. The null hypothesis that the relationship in the total population between M change and therapeutic change is a chance relationship can not be rejected.

The chi square test, rather than the Pearson r, was used, because the distribution of scores for the therapists' ratings of therapeutic change is highly kewed. The range of scores is from zero to fifteen, with the mode at zero and the mean at 2.8. The distribution of M-set change scores (for the same patients) appears to approximate a normal distribution. It has a range of scores from minus three to plus 17, with the mean at plus 4.7 and a standard deviation of 4.88. As the average scores for some therapists' rating of therapeutic change

for their patients differed significantly from the average for the entire group of patients (see page 69), we were forced to conclude that at least one of the therapists' ratings came from a different universe of ratings from that of the other therapists' ratings.

On the assumption that the difference obtained is a function of the therapists' needs, rather than a reflection of real differences in therapeutic improvement, due to systematic differences in other relevent factors, compulsivity, resistance to treatment, motivation, dependency needs, severity of illness, duration of treatment or therapeutic skill, each therapists' ratings were converted into sigma scores, such that the distribution of each therapists' scores would have a mean of zero and a sigma of one.

This statistical procedure removed the effect of the therapist difference entirely, both their individual differences in spread of scores and any deviations from the general mean ratings.

These normalized scores for therapeutic change were then correlated with the scores for M-set shift. The Pearson r is -.10, which is not a statistically significant correlation for this small sample. Thus, with the ratings of the therapist made statistically comparable, we still cannot demonstrate any relationship between change in M due to M set and change in therapy as rated by the patients' therapists.

An attempt was made to see whether the therapists' ratings of therapeutic improvement is related to the length of treatment for each patient. It was believed that there should be a positive relationship between these two variables and that if such a relationship existed in our data, it would lend some support to the validity of the therapists' ratings. For this test, all patients who had less than 10 hours of treatment were excluded. The Pearson r for the 14 remaining cases is +.43. This is significant at the 10 per cent level of confidence. Although this coefficient does not give much statistical support, it is of the magnitude we would expect to find and strengthens our feeling of confidence in the therapist's judgments. We believe the therapists' judgments have validity over and above their bias.

This finding of no relationship between M set change and therapeutic change was not accepted as final. Attempts were made to determine whether some relationship exists between the capacity to shift in the test situation and

therapeutic change, which was not observed in our first test of the hypothesis because of the inadequacy of our criterion measures of shift in M and of change in therapy. Attempts were made to refine these measures and to obtain new measures. Before getting a new measure of change in therapy, refinements of the measure of M change were made, and the data on the M-limits test as a further evidence of capacity to shift in the test situation were studied.

10. Refinement of Measures of M Set Change

One of the examiners had suggested that some of the human movement responses given in the experimental retest did not seem to be "genuine" movement responses. There seemed to be a "forced" and "artificial" quality to the responses. It was thought a refinement of the scoring of M might give a more meaningful and valid measure of shift in perception of human movement.

Three new measures of M change were obtained with the thought that they would reflect a more genuine quality of M response change. In the clinical evaluation of M responses, each response is not given equal weight. The new measures are an attempt to convert some of this qualitative evaluation into quantitative form. The three new measures are:

1. Change in M with plus form quality (M plus change)
2. Change in "weighted M" (Mw change)
3. Change in "Good M" (Mg change)

The weighted M score is obtained by giving a value of one for each of the following characteristics in each M response. Each of these characteristics is generally considered to be an indication of a "good" or genuine M.

a) Form quality -- plus

b) Location -- whole or normal detail

c) Content -- a whole human figure (no weight for human detail, abstraction, animal, or mythological content)

d) Sequence -- M as the first response to a blot or as the response immediately following a popular response, as the first response to a blot.

The good M score is obtained by counting only those M responses which are of plus form quality and located in the whole blot or a normal detail and have a whole human content.

Rank order correlation were computed to determine the relationship between the three refined measures of M change and the simple count of M change in which no account is taken of form quality, location, content or sequence. The following rank order correlations were obtained for all 26 cases in the experimental group:

1. M change and M- change, +.875
2. M change and Mw change, +.969
3. M change and Mg change, +.856

As was expected, the various measures are all positively related. The relationship between M- change and Mw change with M change is so close as to suggest that they are actually

not different measures.

There is a possibility that Mg change may have some independence of M change and can be used as another measure of M set shift.

However, when the three refined measures of M change were correlated with M change for a more limited sample, the 12 subjects in the experimental group who were given the M-limits tests, the following quite different coefficients were obtained:

1. M change and M change, -.45
2. M change and Mw change, -1.00
3. M change and Mg change, -.96

It was concluded that for the M limits group of subjects, only M change is significantly different from M change and can be used as another measure of M change.

11. The M Limits Test

The amount of increase in M in the second test given under the experimental conditions compared with the amount of M in the first test, was conceived as our first and most important measure of capacity to shift. The score on the M limits test was conceived as a second measure of capacity to shift.

The data obtained from the M limits tests for the 12 patients in the experimental group who were given the M limits test were studies. An M limits score was obtained by giving a weight of two to each concept on the M limits test which was accepted at Level 1 or 2, and a weight of one for each concept

accepted at Level 3 (see Appendix II).

The M limits score was found to be statistically unrelated to the M set score and to M on the initial Rorschach. The following rank order coefficients were obtained, none of which is significant.

1. M limits score and M set score -.13
2. M limits score and M change between initial test and M set repeat test -.05
3. M limits score and M on initial test -.11

It was shown, that for this sample, the M limits score is independent of the number of M responses on the initial Rorschach test. It was concluded that the M limits test gives a measure of capacity to shift which is independent and unrelated to the capacity to shift score obtained by the change in M between the initial Rorschach and the experimental Rorschach.

The relationship between the M limits score and therapeutic change was then determined. The Pearson correlation between M limits score and the raw score for the therapists' evaluation of therapeutic change proved to be -.04, which is not significantly different from zero. Similarly, after the therapists' ratings were normalized the Pearson correlation between M limits score and therapists' normalized ratings of therapeutic change proved to be -.12, which is also not significantly different from zero. It was concluded that within the limits of this study the M limits test is not related to therapeutic change for the 12 subjects in the

experimental group who took the M limits test.

As the two tests of capacity to shift are independent, it was believed that combining the two measures would give a more adequate measure of this capacity than either one alone. This was done after another criterion of therapeutic improvement was obtained.

12. Differences Between Therapists' Average Ratings of Therapeutic Change

On page 61 it was stated that examination of the rating of therapeutic change showed marked differences in the average score given to therapeutic patients by each therapist . All the therapists were psychoanalytically oriented and had a similar fundamental outlook with respect to their conception of what are the important elements in psychotherapeutic improvement. Further, the evaluations were made in response to the same specific questions on the form for rating therapeutic change which was devised with the cooperation of the therapists. This similarity in theoretical viewpoint and the use of the same rating form should make for comparability of judgments. Nevertheless, there were probably a number of reasons why therapists showed marked variations in their judgments of therapeutic improvement in their patients. It is probable that the therapists were using different bases for their judgments. It was suggested further that certain needs of the therapists were influencing their evaluation of the progress made by their patients and hence contaminating the results. As the assignment of patients was made primarily on the basis of the

position of an applicant on the waiting list when a therapist had time available for a new patient, rather than on the basis of an estimation of degree of pathology or prognosis, etc., and as the therapists involved are of roughly equal training and experience in psychotherapy, it was believed that there should be no significant difference in the average improvement in groups of patients assigned to each therapist. We tested the hypothesis that there is no statistically significant difference in the average improvement of each therapists' patients from the average improvement in the entire group of patients.

Table 15 shows the computation of the X^2 test for this hypothesis. In this table the observed frequency (fo) is the average score for each therapists' patients obtained from the therapists' "Rating of Therapeutic Change." The theoretical frequency (ft) is the average for the entire group of 17 patients reported in the table. All cases are included in which the therapist had three or more patients in the experimental group.

TABLE 15

CHI SQUARE TEST FOR SIGNIFICANCE OF DIFFERENCES BETWEEN AVERAGE THERAPISTS' RATINGS OF THERAPEUTIC CHANGE

Therapists	N	fo	ft	fo - ft	$(fo - ft)^2$	$\frac{(fo - ft)^2}{ft}$
1	4	12.25	5.59	6.66	44.36	7.94
2	4	6.00	5.59	0.41	0.17	0.03
3	3	3.33	5.59	2.26	5.11	0.91

Table 15 (continued

Therapists	N	fo	ft	fo - ft	$(fo - ft)^2$	$\frac{(fo - ft)^2}{ft}$
4	3	3.00	5.59	2.59	6.71	1.21
5	3	0.67	5.59	4.91	24.11	$X^2 = \frac{4.31}{14.39}$

With four degrees of freedom a X^2 of 14.39 is significant at the 1% level. The null hypothesis may be confidently rejected. It is suggested that the significant differences in ratings given by the various therapists may be a reflection of the expression of certain needs of the therapist. These differences will tend to minimize any relationship that may actually exist between the experimental variables and the therapists' rating of therapeutic change as the criterion of therapeutic change. The desirability of getting another criterion measure was demonstrated.

13. Consultant's Evaluation of Therapeutic Change

It was believed that an improved measure of therapeutic change could be obtained from the psychiatrist-in-chief, who is the consultant to all of the therapists. The consultant's cooperation was solicited. He proved to be extremely cooperative and intensely interested in this research. He was asked to rank the 10 patients in the experimental group who were in treatment at the time, after his individual consultation with each of the therapists involved. He was asked to make three independent rankings for each patient: a) overall judgment of fundamental change in personality; b) judgment of amount of

symptomatic improvement, and c) improvement in terms of his estimate of the original potentiality for change; in other words, improvement relative to the original prognosis. The consultant undertook the task quite seriously and despite difficulties involved made the rankings.

The first step in the analysis of the consultant's ranking was to see whether he was able to differentiate the patients on the three criteria of improvement. The second step was to see whether his ranking differed from that of the therapists' rating of therapeutic improvement. Finally, the third step was to determine the relationship between the experimental variables and this new criterion.

The consultant was unable to differentiate the 10 subjects on the three scales: A) overall improvement, B) symptomatic improvement and C) overall imrpovement in terms of expected improvement. The following rank order correlations were obtained:

A and B, + .95
A and C, + .97
B and C, + 1.00

Inasmuch as there was practically no difference in the rankings, the consultant's rankings were combined and used as one criterion measure.

The consultant's ranking of therapeutic improvement was then compared with the therapists' ratings, after the therapists' ratings were converted into ranks. For the 10 patients ranked by the consultant, the rank order correlation with therapists' ratings was + .58. Because of the small size

of the sample the correlation of +.58 is not quite significant at the 5% level. If the N were a little larger and the same correlation were obtained it would be more significant statistically. It is suggested that the consultant's ratings and the therapists' ratings have some validity. The therapists and the consultant see some parts of the changes taking place in the patient. The small size of the sample limits the decisiveness with which any conclusions may be drawn.

The first application of the consultant's ranking to the experimental variable M change produces a rank order correlation of -.38 (N=10). (The therapist's ratings correlated with M change for the same 10 patients is -.05.) Although neither of these coefficients is statistically significant, there is a suggestion of a negative relationship between the consultant's rating of improvement and M change. This suggested relationship was investigated further by using the M limits test as an additional measure of shift.

Earlier in this chapter it was demonstrated that for the 12 patients in the experimental group who had been given the M limits test, the M limits score is not statistically related to M set change, M+ change or to M on the initial test.

It was concluded that the M limits test gives a measure which is independent of the M set score. As the two tests are independent, and both may be regarded as measures of capacity to shift, it would be expected that combining the two measures would give a more adequate measure of this

personality variable than either one alone. This was done.

The <u>therapists'</u> rating of therapeutic change correlated with the combined measure of M change plus M limits score -.73. (Rank order, N = 8).

The <u>consultant's</u> rating of therapeutic change correlated with the combined measure of shift -.74 (rank order, N = 8).

These results demonstrated that with the combined measure of shift, there is a negative relationship between capacity to shift and each of two different criteria of therapeutic improvement.

A summary and discussion of the findings will be found in Chapter V. It was felt that the failure to obtain a positive relationship between M set shift and psychotherapeutic change might be due, in part, to a tendency on the part of the therapists to see improvement more easily in patients who were more ill to begin with than in patients less severly disturbed. An attempt was made to determine if this were actually the case. This investigation, which was not planned for in the original design of the experiment, is reported in the following section.

Our preliminary analysis of the data to this point suggested that a negative relationship existed between the combined measure of shift in the test situation and change in therapy. In our thinking concerning this relationship and from reading Fisher's monograph (8), we developed the hypothesis that test shift may actually be a reflection of

manifest dependency needs, rather than an indication of healthy type of personality flexibility. The attempt made to determine if there was any validity to this hypothesis is reported in section 15 of this chapter.

14. Rorschach Neurotic Signs and Therapeutic Change

We stated in the last section that the failure to obtain any relationship between M set change and therapeutic change may in part result from a tendency for the therapists to see improvement more easily in patients who were more ill to begin with than to see any improvement in patients initially less severely disturbed. In this section we report on a study designed to test this concept.

Two methods of obtaining measures of degree of maladjustment from the initial Rorschach were suggested. The Monroe check list (20) was rejected in favor of the Miale and Harrower-Erickson neurotic signs (17), because the use of the Monroe check list would have presented a formidable task. The use of the Monroe check list would require the rescoring of all the initial Rorschach protocols according to the Klopfer scoring system and making 28 different judgments for each individual Rorschach record. The Miale and Harrower-Erickson neurotic signs are much easier to determine and were thought to be adequate for our purpose.

The Miale and Harrower-Erickson neurotic signs are: 1) 25 R or less, 2) Not over 1M, 3) Refusal to respond on one or more cards, 4) Over 50% F responses, 5) Over 50% A plus Ad responses, 6) Not over 1FC, 7) Color Shock, 8) Shading Shock.

Each sign, when present, is given a score of one toward the individual's Neurotic Score. We used the first six of the eight signs, because, 1) some of our protocols were not complete enough to permit the accurate determination of the presence of color and shading shock, and 2) the scoring of color and shading shock involves some subjective judgment, while all the other signs are determined objectively once the protocol has been scored. As previously reported, our scoring was done with adequate safeguards to assure the elimination of bias in scoring.

The neurotic score obtained and the mean therapeutic change score for the patients with each neurotic score is presented in Table 16.

TABLE 16

NEUROTIC SCORES (6 of 8 MAILE AND HARROWER-ERICKSON NEUROTIC SIGNS) AND SCORES FOR THERAPISTS' RATINGS OF THERAPEUTIC CHANGE

N	Neurotic Score	Therapeutic Therapeutic Change Mean Score
3	0	8.33
2	1	3.00
3	2	6.00
8	3	5.00
2	4	2.50
2	5	1.00

Statistical analysis revealed no significant relationship between the number of neurotic signs and therapeutic change as judged by the therapists. (See Table 17.)

TABLE 17

RELATIONSHIP BETWEEN NUMBER OF NEUROTIC SIGNS AND THERAPEUTIC CHANGE

	Therapeutic Change Score		
Neurotic Signs	Low (0 to 3)	High (4 to 15)	
High (3 to 5)	7	5	12
Low (0 to 2)	3	4	8
	10	10	

chi square = .734
P = 30 - 50

The data suggests the conclusions, that for the entire group of 20 experimental subjects who entered treatment there is no relationship between severity of neurosis as judged from the initial Rorschach and change in therapy as judged by the therapists. The further conclusion would follow that the failure to obtain any significant relationship between M set change and therapeutic change is not due to a tendency for the therapists to see improvement more easily in patients who were more ill before treatment compared with seeing any improvement in patients initially less severely disturbed. It might be implied, that the therapist's judgment of therapeutic change was made independently of the severity of the patients' neurosis.

It was believed that the Rorschach neurotic signs score might give additional information about the nature of M set shift and suggest something about the validity of the psychiatrists' ratings of initial illness. A Pearson r of

+.14 was obtained between psychiatrists' ratings of initial illness and Rorschach neurotic score for the 26 patients in the experimental group. The correlation is not statistically significant.

A Pearson r of -.32 was obtained between Rorschach neurotic score and M set change. This correlation is not statistically significant. The chi square test for this relationship is also not significant (see Table 18).

TABLE 18

RELATIONSHIP BETWEEN NUMBER OF NEUROTIC SIGNS AND M SET CHANGE

	M Set Change		
Neurotic Signs	Low (-3 to +2)	High (+3 to +17)	
High (3 to 6)	9	6	15
Low (0 to 2)	4	7	11
	13	13	

chi square = 1.72
P = 10 - 20

Although both the r and the chi square test are not statistically significant, they suggest a tendency for the more severely neurotic patients to shift less in M than the less severely neurotic. This finding tends to support the original major hypothesis that M set shift is an index of mental health.

15. Dependency Needs and Shift in the Rorschach

It has been shown that shift in the Rorschach is negatively related to change in psychotherapy. This negative

relationship is found when shift is measured by a combined score of M change between a standard Rorschach and an M set retest, and an M limits test, and when change in psychotherapy is evaluated either by the therapists' ratings or by a consultant's rating. This test shift, or instability, is negatively related to the capacity for making a "deeper" change which is necessary for progress in psychotherapy. It is suggested that test instability may be a phenomenan related to the manifest expression of dependency needs and may be understood in terms of prestige suggestion, superficial cooperation with and need to please the examiner. We may think of test shift as being positively related to adjustment but negatively related to mental health.

This conception of the nature of test shift would be supported if it could be demonstrated that there is a positive relationship between test shift and manifest dependency needs. In this portion of the experiment an attempt was made to study this relationship.

To our knowledge there is no published scale for determining the extent of dependency needs. However, as a part of his doctoral research[1], Stanley Schneider had constructed scales for a number of variables, including dependency (Appendix III). The scale includes a place for the rater to indicate whether each of the statements checked as applying to the patient's relationship to the therapist, was based on manifest behavior or was implied from latent behavior.

---

[1] In process of completion at the University of Michigan, 1952.

The therapists completed a scale for each of the 16 patients in the experimental group who were in treatment at the time. Three scores were derived from these scales; a) overall dependency; b) manifest dependency; and c) latent dependency. The overall dependency score is the sum of the statements checked as applying to the patient, regardless of whether manifest or latent or both. The manifest dependency score is the sum of the statements checked as being manifest. The latent dependency score is the sum of the statements checked as being latent.

The following rank order correlation coefficients were obtained for all 16 cases:

1) Overall dependency and manifest dependency, +.91
2) Overall dependency and latent dependency, -.41
3) Manifest dependency and latent dependency, -.34

It is concluded that on this scale, overall and manifest dependency are highly related and are not independent measures. There are tendencies not statistically significant, for overall dependency to be negatively related to latent dependency and for manifest dependency to be negatively related to latent dependency.

Correlating the three dependency measures with shift in M(M_1 - M_2) for all 16 subjects in the experimental group the following rank order coefficients were obtained:

1) Overall dependency and M set shift +.15
2) Manifest dependency and M set shift +.27
3) Latent dependency and M set shift -.38

None of these correlations is statistically significant. However, they suggest that manifest dependency is positively related to shift and latent dependency is negatively related to shift. In other words, patients whose dependency needs are so great that they are expressed openly and directly tend to do as the examiner requests in the M set retest. Patients whose dependency is latent, who for whatever reason it may be, do not express their dependency needs openly and directly, tend not to do as the examiner requests in the M set retest.

When the dependency scales were applied to the nine patients who took the M limits test and correlated with the measure of combined shift (M change M limits), the following coefficients were obtained.

1) Overall dependency and combined shift .05
2) Manifest dependency and combined shift .36
3) Latent dependency and combined shift .63

Again we see that the manifest expression of dependency needs during the therapeutic sessions tends to be positively related to test shift, in this case to the combined measure of test shift and that the latent expression of dependency needs in treatment is negatively related to test shift. The small number of cases limits the confidence with which conclusions may be made. Nevertheless, there is a tendency in our findings which suggests that those individuals who accept an externally given set in the Rorschach and who also respond positively to a "testing the limits" procedure, may be expressing their strong dependency needs and actually be making a

bid for the examiner's approval. This concept deserves rigorous investigation through a study specifically designed to test it.

CHAPTER V

SUMMARY OF FINDINGS AND DISCUSSION

Now we will summarize our argument with respect to the principal hypothesis. The principal hypothesis is that test shift is an index of mental health and is positively related to change in therapy. We have found that neurotics as a group are able to shift in the experimental situation with ease and to an extent that produces significant alteration in the M:C ratio. But what does this capacity to shift mean in term of personality dynamics and in terms of capacity to change in psychotherapy? Is the capacity to shift in the test situation an index of mental health, of a healthy flexibility and integrative capacity of the ego? So far, the evidence does not support this hypothesis.

The data concerning the similarities and differences between the normals and the neurotics, with respect to M change, are not entirely clear as to their meaning. The neurotics are lower in mean initial M and increase less in terms of raw M score, than the normals. However, on a percentage basis, the neurotics increase fully as much as the normals. Further, the proportion of the neurotics and normals who change in direction of M:C ratio is comparable. If capacity to shift in the test situation is an index of mental health, neurotics should shift less readily than normals.

This finding does not support the original hypothesis. This result would suggest that capacity to shift is not related to mental health.

We found that the five initial Rorschach determinants and relationships between determinants that characterized Milton's shift-most normal group, did not differentiate between our neurotic shift-most and shift-least groups. In fact, two of his three factors found to characterize his shift-most group were found to characterize our shift-least group. This includes the factor (W minus M of ≠ 1 to ≠ 7) which was the only single item which was significant for the normals. Milton states "A W minus M score of between ≠ 1 and ≠ 7 indicates adequate ego strength (in our conception of the term)." If Milton's conception is correct, then we must conclude, that at least for our sample of neurotics, patients with adequate ego strength tend to shift less in the experimental situation than patients who lack adequate ego strength. This is the reverse of our original hypothesis.

We found no significant difference in M set shift for patients low in initial M compared with patients high in initial M. The high initial M group would be expected to increase M more than the low initial M group if test shift is conceived as an index of mental health and initial M is accepted as a sign of ego strength, either as a measure of current intellectual functioning, as most prominent Rorschach writers affirm, and as is supported by considerable research; or if M represents a stabilizing influence against the acting

out or direct expression of impulses, as Hutt believes. This finding was not altered by modifying the measures of M set change in an attempt to "purify" the M score in terms of measures of "good M", "weighted M" or "M ≠".

We found, 1) patients with low initial M (0 to 2) give moderate increase in M (≠ 1 to ≠ 7); 2) high initial M patients fall into two distinct groups; a) those who show marked increase in M (≠ 8 to ≠ 17) and b) those who show no increase in M (-3 to 0). This suggests that the presence of a good deal of M in a Rorschach record has different meanings for different people. Our exploratory study tends to support the original hypothesis of M shift capacity as an index of mental health, as those high initial M patients who increase M show a richer personality structure on the original Rorschach than those who do not shift.

The finding of no statistical relationship between M set shift and the psychiatrists' overall initial rating of degree of illness, prognosis and compulsivity, does not confirm the original hypothesis. The tendency for a negative relationship between M set shift and psychiatrists' judgment of severity of illness, tends to support the reverse of the original hypothesis. However, the negative relationship between M shift and psychiatrists' judgment of compulsivity tends to support the original hypothesis.

The finding of no statistical relationship between Neurotic Index (number of "neurotic signs" in the initial Rorschach tests) and M-set shift, does not support the original

hypothesis. Neurotic subjects who presumably are less severely neurotic according to the initial Rorschach, do not increase their productivity in M in response to the external set, more than subjects who are presumably more severely neurotic.

Now we will bring together the evidence on the nature and meaning of M set shift, before considering the data on therapeutic change.

For support of the original hypothesis of M set as an index of mental health:

1. Neurotics increase less in absolute number of M than the normals.
2. High initial M neurotics who increase M show richer personality structure on the initial Rorschach, than high initial M patients who do not increase M on the retest.
3. Neurotics judged more compulsive by intake psychiatrists shift less than patients judged less compulsive.

Against support of the original hypothesis of M change as an index of mental health:

1. Neurotics increase as much in M as normals in terms of percentage of initial M score.
2. The proportion of neurotics and normals who change their M:C ratio is comparable.
3. The absence of a simple relationship between number of initial M responses and M set change.
4. The absence of a relationship between Neurotic Index (based on the initial Rorschach) and M set change.
5. Two of the three factors in Milton's stability score which characterize his shift-most group, characterize our shift-least

group.

6. The absence of a relationship between M shift and psychiatrists' original overall rating, which includes: severity of illness, prognosis, and compulsivity.

For support of the reverse of the original hypothesis:

1. The W-M of plus one to plus seven which characterizes our shift least group and which is interpreted by Milton as an indicator of adequate ego strength.

2. The tendency for a negative relationship between M set shift and psychiatrists' judgment of severity of initial illness.

However, conclusions based on comparisons between Milton's normal group and our neurotic group should be interpreted with caution and not generalized without additional evidence based on further study with other groups of normals and neurotics and with better matching of these groups. If possible the motivation for the two groups should be equated. Our neurotics were not matched with the normals, although effort was made to assure some comparability in age and educational level. There were also differences in the examiners. The normals were examined by a large number of rather inexperienced examiners, while most of the neurotics were examined by experienced clinicians. It may well be that the experienced examiners were able to elicit more cooperation from the subjects. The normals were cooperative student volunteers who were probably not emotionally involved in the task. The neurotics were applicants for treatment who were probably much more ego

involved in the task and therefore more strongly motivated to cooperate with the examiners. It is reasonable to believe that the neurotic group produced results more nearly approximating their capacity as a result of their having been tested by experienced examiners and having been more ego involved in the task. Therefore, the approximately equal shift found in the two groups may not reflect approximately equal capacity to shift. The hypothesized greater capacity for the normals to shift, compared to the neurotics, may still be a valid concept, despite the failure of this experiment to demonstrate such a difference.

Although the evidence is not clear cut, taken all together, it does not support the original hypothesis of test shift as an index of mental health. The tentative conclusion at this point is that test shift is not related to mental health or ego strength.

Now we will continue with the summary of the argument, bringing in the findings with respect to therapeutic change.

The <u>therapists'</u> ratings of therapeutic change do not support the original hypothesis. Neither the raw scores nor the normalized scores of therapeutic change are statistically related to M set shift.

The <u>consultant's</u> rankings of therapeutic change suggest the reverse of the original hypothesis may be true. We obtained a rho of -.36 between M set shift and consultant's ranking of 10 patients on therapeutic improvement.

The data obtained by correlating the ranking of

patients on the combined score of M set shift and the M limits score with therapeutic change tends to support the reverse of the original hypothesis. The rank order correlations for combined shift and ratings of therapeutic change are: 1) -.73 with therapists' ratings (N = 8); 2) -.74 with consultant's ranking (N = 6).

Our conclusion is that, within the limits of this experiment, M set shift is definitely not positively related to therapeutic change. There is some indication that it may be negatively related, but this latter conclusion is based on only 10 cases and the negative correlation obtained has little statistical significance. The conclusion which has the greatest support is that of no relationship between M set shift and behavioral change as judged from therapeutic change. This conclusion is congruent with our previous findings of the relationship of M set shift to initial Rorschach determinants, to the psychiatrists' initial ratings, and to the relationship between our findings and those of Milton on normal subjects.

Our data further indicate that the M limits score and the M set change score are independent of each other and of therapeutic change. However, when the M limits score and the M set change score are combined, the combined score is negatively related to therapeutic change. It is suggested that neurotic patients who accept both the externally imposed M set and accept the suggestions offered by the examiner in the M limits test, have less ego strength than patients who do not accept both situations. They may be more dependent indi-

viduals whose dependency is manifest. They may have less confidence in themselves and change with every alteration in the environment. They show the increased suggestibility people generally show when their own confidence level is low. On the other hand, they must have the capacity to see human movement, to find M in the cards, as the M is not pointed out by the examiner in the M set retest and is shown specifically in the M limits test only on the third level, which is given little weight in the M limits score. We would hypothesize that the patients who do not shift in the M set retest or the M limits test do not necessarily lack capacity to shift. They may have less need to please the examiner and more determination to stick to their own initial responses. Perhaps this may be regarded as parallel to a latent learning situation. The more secure, more adequate, less dependent patients had less internal motivation to perform for the examiner. Perhaps, given adequate motivation, they would achieve the same level or exceed the more dependent patients who were more motivated to perform and whose learning was manifest.

Our findings may also be examined in terms of situation versus personality dominance. As Brownfain (5) expresses it: "The individual with high self-esteem is less likely to see himself as dominated by a situation. His self-concept is relatively stable because he is sufficiently the master of his environment that a radical restructuring of the self-concept need not be the consequence of changing situations." The po-

sition that Brownfain postulates with respect to the self-concept may well apply to other concepts, including concepts as to what the Rorschach ink blots suggest to the subject.

We may also think in terms of "ego" rigidity and "pripheral" rigidity (Fisher, 8), and in terms of response stability versus personality stability. It may be fruitful to think in terms of rigidity behavior as correlated with a series of "defense levels" in the personality structures.

At this stage in our speculation it was felt that it would be fruitful to gather some data on the relationship between manifest and latent expressions of dependency needs and test shift. It was also believed that the failure to obtain any relationship between M set change and therapeutic change may in part result from a tendency for the therapists to see improvement more easily in patients who were more ill to begin with, compared with seeing any improvement in patients initially less severly disturbed.

Two additional studies were then carried out in an attempt to clarify the meaning of our findings. The first study was on the Rorschach signs of neurosis and therapeutic change. The second study was on the relationship between dependency needs and shift in the Rorschach.

We found no relationship between severity of neurosis as measured by the number of neurotic signs in the initial Rorschach and therapeutic change as judged by the therapists. This finding suggests that the therapists' judgments of therapeutic change were not seriously attenuated by a tendency to

see improvement more easily in the more ill patients than in the less ill patients.

The hypotheses tested in the dependency study were, 1) the manifest expression of dependency needs is positively related to shift in the Rorschach, and 2) latent dependency is negatively related to test shift. Our findings tend to support both hypotheses. For the 16 patients in the experimental group who were still in treatment at the time, we obtained a rank order correlation between manifest dependency and M set shift of ρ .27. For the same patients, the rank order correlation obtained between latent dependency and M set shift was -.38. These correlations, while not statistically significant, are in the predicted direction. The hypotheses were more strongly supported when the dependency scores were correlated with the measure of combined shift (M set change plus M limits score) for the nine patients still in treatment who took the M limits test. We obtained a rank order correlation between manifest dependency and combined shift of ρ .36. For the same patients we obtained a rank order correlation between latent dependency and combined shift of -.63.

In other words, the measure of test shift which showed the highest negative relationship with therapeutic change (the combined shift score) is the same measure which shows the highest positive relationship with manifest dependency and the highest negative relationship with latent dependency. These findings lend considerable support to our

speculation that in this study, shift in the test situation is a function of manifest dependency needs.

We cannot say that test shift is an index of mental health, nor can we say that test shift is unrelated to mental health. Rather, we must revise our concept of the index. The important question is not that of more or less shift. The significant question is how appropriate is the shift. In this study it is difficult to determine what the most appropriate level of shift would be. In the concluding chapter some suggestions for research in this area are presented.

We now think that the optimal integrity of the personality envolves a sufficient self-awareness and acceptance of self to appropriately resist or adapt to changes in the environment. In other words, the healthy individual is one who is able to shift with changes in the environment as he chooses to do in accordance with his own needs. In the healthy personality, adaptation is reality oriented in terms of ego needs. It is capable of those kinds of shift which do not disturb its own integrity.

The basically insecure individual, e.g. the authoritarian personality, has to succumb to every change in the situation irrespective of ego needs and its own integrity. It does not shift appropriately but wallows in the stimulus. Or else, the basically insecure individual may react in exactly the opposite, but equally inappropriate, manner. It may never shift under any circumstances. An example would be the psychotic paranoid individual with well systematized delu-

sions upon whom reality doesn't impinge.

CHAPTER VI

GENERAL SUMMARY AND CONCLUSIONS

In this, the final chapter, will be found a brief summary of the experiment, a statement of the conclusions and implications, and suggestions for further research.

This study is concerned with predicting therapeutically induced changes in behavior from variables especially derived from Rorschach's test. Therapy is regarded as a prolonged interpersonal relationship which provides a maximal condition for the induction and observation of behavioral and perceptual change in the patient.

The experimental condition consisted of re-administration of the Rorschach test with altered instructions, requesting the subjects to modify their performance on the human movement test factor. An additional experimental condition consisted of an "M limits" test designed to establish the full extent of the subjects capacity to utilize this significant Rorschach variable, human movement.

The subjects were 40 young adult psychoneurotics who had been accepted for treatment at a public mental health clinic at the time of testing. Each subject was given the Rorschach twice within an interval of approximately two weeks, the first time under conditions of standard instructions. In the second testing, the 14 control subjects were again tested

under standard instructions, while the 26 experimental subjects were each examined under the altered instructions. The M limits test was given immediately after the second testing to a portion of experimental groups. A portion of the control group was also given the M limits test and was used as an experimental group for this share of the experiment.

The intake psychiatrist evaluated the subjects severity of illness, prognosis, and compulsivity on an initial rating form. Approximately six months after treatment had begun, or after termination if treatment was discontinued within six months, the therapist evaluated the changes resulting from treatment on a Rating of Therapeutic Change form devised for this study. Due to marked differences obtained between the average ratings of therapeutic change made by the several therapists, a second criterion measure of therapeutic change was obtained from the psychiatric consultant.

A study of the relationship between manifest and latent dependency needs and shift in the test situation, which was not included in the original design of the experiment, was undertaken to assist in evaluating some of the findings of the overall experiment.

The Rorschach testing was done by examiners other than the experimenter and adequate safeguards were taken to assure elimination of bias in scoring, which was done according to Beck's scoring system.

In deriving conclusions and implications from this

study attention must be drawn to the nature of our sample of neurotic subjects and to the nature of the sample of normals with which our findings were compared. It was noted, for example, that with respect to the particular variable investigated in this study, the M factor, both our neurotics and Milton's normals are more productive on the average than is generally reported in the literature. There are also differences in a number of factors which may be sources of variance in the Rorschach protocols. Some of these could not be controlled because of the fact that the normal sample consisted of college students tested in an academic setting by inexperienced graduate students who were getting experience in testing, while the neurotic sample consisted of young adults who were disturbed to a degree that brought them to seek professional assistance with their emotional problems at a public clinic. The neurotics took the test after intake interviews with a psychiatrist and a psychiatric social worker, while waiting to be assigned to a therapist, and were examined by experienced clinicians. Miller (18) has summarized the findings reported in the literature pertinent to any interpersonal relationship, including diagnostic testing and therapy. These five categories of variables are: 1) setting, 2) task, 3) examiner's social stimulus value and character structure, 4) subject's character structure, and 5) relationship between examiner and subject. In each of these areas there are variations within each group and probably significant differences between the normal and neurotic groups. Nevertheless, the portion of this experiment that

deals with the initial and repeat Rorschach tests was designed to be comparable with the procedure used for the normals. Of Miller's five categories of variables that seem pertinent to any interpersonal situation, only the "task" could be made reasonably comparable for the two groups.

Without forgetting the preceeding considerations, the following conclusions derived from this study, may be applicable to psychoneurotics generally.

1. The Human Movement factor in the Rorschach is unstable. Implications: a) The score on a given Rorschach variable is a product of many factors both internal and external. One is not justified in saying that an obtained score is the score for an individual. For more adequate interpretation many other factors than those which appear in a Rorschach record should be considered.

b) The instability of the Rorschach factors does not mean that the Rorschach technique is unreliable. Rather, it may be an indication that it is a sensitive instrument. When interpretations are made, one must appraise both the individual and the total test situation.

c) It is inferred that as the M factor can so easily be modified, other factors which are presumed to be less unconsciously determined and more subject to voluntary change, could be even more easily modified by the conscious effort of the subject.

d) It is inferred further, that as the M factor can be modified by an external set, any internal set brought by

the subject to the test situation would also affect his production on the Rorschach.

2. There is no simple relationship between the number of M's on the initial test and M set shift. Subjects who give no or few M responses in the standard test do not show less increase in the experimental retest than subjects who give many M responses on the standard test.

Implication: The score obtained on a Rorschach is not a valid index of the capacity of the subject. Rorschach's statement that only "Subjects who in the ordinary experiment give numerous M's can increase the number of kinaesthetic responses with relative ease", is not supported.

3. The experience balance is unstable.

Implications: a) Rorschach's belief that "the absolute number of M and C responses changes but the proportion between them changes little or not at all", is not supported.

b) In interpreting the E. B., the individual and the entire setting must be evaluated. The changes in E. B. resulting from the external set would produce very different personality descriptions for most of the subjects if the set were not considered in the evaluation.

4. Neurotics shift in M and experience balance about as readily as normals. Implication: Some doubt is cast upon "capacity to shift in M" as an indication of capacity for behavioral change or an index of mental health. If our original concept were correct, a group of normals would be expected to shift significantly more than a group of neurotics.

The following conclusions are drawn from our findings with respect to therapeutic change as evaluated by the therapists and by the consultant.

5. The initial M score is not significantly related to therapeutic change.

6. The M set shift score is not directly related to therapeutic change.

7. The M limits score is not directly related to therapeutic change.

8. The combined measure of shift is negatively related to therapeutic change.

The following conclusions are drawn from our findings with respect to the relationship between dependency needs and test shift.

9. The extent of manifest dependency needs is positively related to M set shift and to the combined measure of shift.

10. The extent of latent dependency needs is negatively related to M set change and to the combined measure of shift. Implications: Our conception of shift in the test situation as an indication of a healthy flexibility in the personality structure must be revised. The relationship between test shift and the capacity for making more significant changes in behavior and in the perception of the self and others, is a more complex one. The significant question is not how much or how little does the individual alter his response. We believe that the healthy individual is one who

is able to shift or not shift with changes in the environment as he chooses to do in accordance with his needs. The healthy personality is capable of those shifts which do not disturb its own integrity. The basically insecure individual goes to one extreme or another. He either has to succumb to every change in the situation irrespective of his own integrity or he may never shift under any circumstances.

In brief, we have found that neurotics are able to alter their perception in a test situation with ease about as much as normals. But this capacity to shift is not a reflection of a healthy flexibility. Rather, it is suggested that following the instructions of the examiner in both the M set retest and in the M limits test is an expression, in part, of strong dependency needs. However, this implication is difficult to support at this point since the normal and neurotic samples were not matched for relevant factors. Also, we have some evidence within this study that shift in M may be a reflection of a healthy flexibility. This evidence is the negative relationship found between psychiatrists' rating of compulsivity and M set shift. This study should not be taken to mean definitely that M set shift is not related to mental health. Before we can generalize the results of this study, these and other interrelationships would have to be studied intensively to develop a more appropriate theoretical framework.

This research, like most research, grows out of past investigations and builds on to it. This study, in turn,

raises more questions than it answers. In order to better understand some of the findings and formulations offered in this paper, the following areas of study are suggested as worthy of research effort.

1. A repetition of this study at another clinic with more training given to the therapists to obtain more valid and comparable ratings of therapeutic change. The use of criteria of personality change based on objective indices of change in the life behavior of the patients, outside the treatment situation, would be desirable.

2. A study of other Rorschach variables under conditions similar to the conditions of this experiment.

3. A study of the applicability of our findings to other nosological groups, e.g. a psychotic group.

4. A study of the stability of M when the change in set is induced by non-verbal means, e.g. frustration.

5. A carefully worked out study of the M limits test in relation to therapy, with a larger sample.

6. A study of the persistence of the experimentally induced change by the use of a third Rorschach after an extended period in treatment for one group and after the same period of time without treatment for another group.

7. A study relating test shift and therapeutic change to the internal interrelationships between compulsivity, dependency needs and numerous other personality variables.

8. A detailed study of the Rorschachs of patients who improved in treatment compared with patients who did not

improve. In such a study the Rorschach test situation should be completely reported, by the use of recordings and detailed report by the examiner of non-verbal behavioral manifestations, to permit the evaluation of elements in the test situation which do not appear in the psychogram.

9. A study of capacity to shift using a role playing technique; e.g., using instructions such as "Can you for the sake of the experiment adapt a framework.."

10. A study of shift with respect to ambiguous stimuli compared with shift when the stimuli are clear cut.

APPENDIX 1

DIRECTIONS FOR EXPERIMENTAL RETEST

"About two weeks ago I gave you the Rorschach examination. Today I want to give you this test again as part of an experiment which we are conducting. When this test is over I shall be glad to discuss the test procedures with you. You will remember that last time I asked you to look at each of the ink-blots and tell me everything you saw, everything they could be. No two people respond to this test in exactly the same way. Now, today, I want you to take this test with a special objective in mind. I want you to tell me everything you see, but in addition I would like to see how many answers you can give in which you see <u>people doing something</u>. You remember this card (showing card III). Some people see two persons here (pointing, quickly to the usual human areas); some see these persons standing, others see them bending, others see them pulling, and still others see them doing other things. (<u>Put card III away</u>) Sometimes no people at all are seen on that card. Well today I would like you to look for people, especially people in motion. As you examine each card, keep in mind that whatever else you may see and tell me about, try to give as many answers as you can about people you can find." <u>At the end</u> of this test, say "That was a very fine job. Thanks for your cooperation in this experiment."

APPENDIX 2

DIRECTIONS FOR "TESTING THE LIMITS" FOR M

Test for each of the responses below that have not been given in the free association or inquiry. Test for each response first at Level I, using such phrases as: "Some people see two people. Can you make them out?" If the subject says "yes, I see it now", conduct the inquiry to find out if he actually perceives it and how he perceives it. If M is not a determinant, ask:

a) "Do you see them in motion or stationary?"

b) "Are they doing something or not doing anything?"

If the subject cannot find it at Level I, try Level II, in which you specify the area in which the percept may be made, saying, for example: "Well, this is where they usually see the people (indicating the area on the card). Can you see it now?" If accepted, conduct the inquiry in the same manner as for Level I, including questions a and b if M is not a determinant.

If the response is still not forthcoming, try Level III, the level of most specific suggestion, saying, for example, "you see, here are the heads and here are the trunks and here the arms (indicating the appropriate areas). Can you see it now?" Again, if accepted, conduct the inquiry as for I or II. If not accepted, drop the matter and record as rejected.

List each response for which a test was made with the appropriate symbols. Where M is not a determinant, record the answers to questions a and b as + or -, with respect to activity in the percept. Thus for the first two figures, the following might appear on the record.

Test the Limits

Fig. I.	Human Form	Level II, F+, a-, b
	Two witches	Level III, M
Fig. II.	Two humans	Rejected

Figure I. Human Form (D3 or D4)
Two witches on a broom (both D2)

Figure II. Two Humans (W or both D)

Figure III. Two figures (D1)
V Two Negroes (D4 and D9)

Figure IV. Man in bearskin coat (W)

Figure V. Person in costume (W)

Figure VII. Human heads or faces (D1 or D1 plus)
V Two dolls with heads joined (W)

Figure IX. > Motorcycle rider (D1)
Witches or goblins (D3)

Figure X. Two fat men (D6)
Two people (D9)

Read each statement carefully. Decide whether or not it applies to _____ in his relationship with you.

	EARLY	MIDDLE	LATER	MANIFEST	LATENT
A1. He needs a good deal of support from me.	____	____	____	____	____
A2. He resents any implication that he "is not his own boss"	____	____	____	____	____
A3. He is afraid to be critical of me.	____	____	____	____	____
A4. He expects me to make decisions for him	____	____	____	____	____
A5. I do not feel dependency is one of his problems	Check if true ____ if untrue ____				
A6. He needs reassurance about my interest in him	____	____	____	____	____
A7. He seems not to really care what I think of him	____	____	____	____	____
A8. He accepts almost everything I say as true and right	____	____	____	____	____
A9. He has threatened to leave treatment, but has never carried this out	____	____	____	____	____
A10. One of his chief concerns is to have me like him	____	____	____	____	____

	EARLY	MIDDLE	LATER	MANIFEST	LATENT
A11. He relies on me to work out his problems	____	____	____	____	____
A12. He tries to involve me with persons outside the treatment (e.g. using your name and position to sanction acts and ideas of his)	____	____	____	____	____
A13. He is uncomfortable when I do little talking in the hour	____	____	____	____	____
A14. He seems to accept the fact that he is a very passive person	____	____	____	____	____
A15. He tries to avoid responsibilities	____	____	____	____	____
A16. He states things here in such a way that he "appears in a good light"	____	____	____	____	____
A17. He acts as if he expects us to manipulate him	____	____	____	____	____
A18. He is uneasy when I do not make interpretations	____	____	____	____	____
A19. He is very demanding of me	____	____	____	____	____
A20. I feel he is very dependent basically, but he tries to deny this to me and to himself	Check if true____ if unture____				

	EARLY	MIDDLE	LATER	MANIFEST	LATENT
A21. He feels that I do not do enough for him	____	____	____	____	____
A22. He frequently asks me for advice	____	____	____	____	____
A23. He must carry out things by himself	____	____	____	____	____
A24. He has difficulty in expressing feelings of needing to be loved and cared for	____	____	____	____	____
A25. He thinks of himself as a person who is usually in control of the situation	____	____	____	____	____
A26. He tends to be deferential and obsequious almost to an extreme	____	____	____	____	____
A27. He cannot express his hostile feelings toward me	____	____	____	____	____
A28. He likes to stay around after his hour is up	____	____	____	____	____
A29. He is very much concerned with what I think of him	____	____	____	____	____
A30. He feels that he is not "good enough" to take up my time in treatment	____	____	____	____	____

	EARLY	MIDDLE	LATER	MANIFEST	LATENT
A31. He must make sure that what he says will be acceptable to me	____	____	____	____	____
A32. He begins the hour with a statement like "I don't know what to talk about today" or "nothing to say"	____	____	____	____	____
A33. He prefers to be questioned by me rather than bring up things spontaneously	____	____	____	____	____
A34. He acted as if I were a "father confessor"	____	____	____	____	____
A35. He acted as though, now that he was in treatment, all of his troubles would be cared for (implies that he would have to do very little by himself)	____	____	____	____	____
A36. He seems to be able to only think well of me	____	____	____	____	____
A37. He admits he needs my help, and at the same time resents me because of this	____	____	____	____	____
A38. He looks upon me as an authority who must be obeyed	____	____	____	____	____

	EARLY	MIDLLE	LATER	MANIFEST	LATENT
A39. He wishes to have my support and affection, but must at the same time rebel against what I do or say	____	____	____	____	____
A40. He thinks of me predominantly as a person who can help him and supply his needs	____	____	____	____	____

APPENDIX 4

1. Footnote to Table 1, page 26.

It was suggested that the significant t's obtained in Table 1 might be spurious, resulting from the skewed distribution of scores and from the fact that the standard deviations tend to vary with the means for each group. It was suggested that the chi square test would be more appropriate. Therefore, the following chi square tests were made.

A. COMPARISON OF HUMAN MOVEMENT RESPONSES ON INITIAL TEST OF CONTROL AND EXPERIMENTAL GROUPS

	Initial M			
	Low	Mod.	High	
Group	(0 to 1)	(2 to 4)	(5 & above)	
Control	6	6	2	14
Experimental	11	9	6	26
	17	15	8	

chi square = 0.345
P = 80 - 90%

B. COMPARISON OF SUM OF COLOR RESPONSES ON INITIAL TEST OF CONTROL AND EXPERIMENTAL GROUPS

	Initial Sum C			
	Low	Mod.	High	
Group	(0 to 1.5)	(2.0 to 4.5)	(5.0 & above)	
Control	7	7	0	14
Experimental	6	10	10	26
	13	17	10	

chi square - 7.721
P - 02 - 05%

C. COMPARISON OF NUMBER OF RESPONSES ON INITIAL TEST OF CONTROL AND EXPERIMENTAL GROUPS

	Initial R			
	Low	Mod.	High	
Group	(under 20)	(20 to 39)	(40 & above)	
Control	5	5	4	14
Experimental	11	5	10	26
	16	10	14	

chi square = 1.346
P = 50 - 70%

It is concluded that the chi square tests give the same results as the t tests. The conclusions drawn from the t and chi square tests would be essentially similiar.

2. Again it was suggested that the significant t obtained might be a spurious one resulting from the skewed distributions. Therefore, the following chi square tests were made.

A. COMPARISON OF HUMAN MOVEMENT RESPONSES ON INITIAL TEST WITH REPEAT TEST FOR CONTROL GROUP

	M Responses			
	Low	Mod.	High	
Test	0 to 1	2	3 or above	
Initial	6	5	3	14
Repeat	7	3	4	14
	13	8	7	

chi square = 0.712
P = 70%

B. COMPARISON OF HUMAN MOVEMENT RESPONSES ON INITIAL TEST FOR EXPERIMENTAL GROUP

	M Responses			
	Low	Mod.	High	
Test	0 to 1	2 to 4	5 or above	
Initial	11	9	6	26
Repeat	3	10	13	26
	14	19	19	

chi square = 7.203
P = 02 - 05%

Again it is concluded that the chi square tests give the same results as the t tests. The conclusions drawn from the t and chi square tests would be essentially similar.

BIBLIOGRAPHY

1. Barrell, R. P. "The Relationship of Various Types of Movement Responses in the Rorschach Test to Personality Trait Ratings," Unpublished Ph.D. dissertation, University of Michigan, 1950.

2. Beck, S. J. Rorschach's Test. Vol. I, New York: Grune and Stratton, 1944.

3. Beck, S. J. Rorschach's Test. Vol. II, New York: Grune and Stratton, 1945.

4. Bell, J. E. Projective Techniques. New York: Longmans, Green, 1948.

5. Brownfain, J. J. "Stability of the Self Concept as a Dimension of Personality." (in press, J. Abnorm. Soc. Psychol.)

6. Burton, Arthur. "The Use of Psychometric and Projective Tests in Clinical Psychology," J. Clin. Psychol., 28, 1949, 451-456.

7. Carp, A. L., and Shauzin, A. R. "The Susceptibility to Falsification of the Rorschach Psychodiagnostic Technique," J. Consult. Psychol., 1950. 14, 230-233.

8. Fisher, Seymour. "Patterns of Personality Rigidity and Some of Their Determinants," Psychol. Monographs, No. 307, Washington D. C. : Amer. Psychol. Assoc. 64, 1950.

9. Gibby, R. "The Influence of Varied Experimental Sets Upon Certain Rorschach Variables: I. Stability of the Intellectual Variables." Unpublished Ph.D. dissertation, University of Michigan, 1949.

10. Hertz, M. R. "Current Problems in Rorschach Theory and Technique," J. Projective Tech., 1951, 15, 307-338.

11. Hutt, M. L., and Shor, J. "Rationale for Routine Rorschach Testing the Limits," Rorsch. Res. Exch., 1946, 10, 70-76.

12. Hutt, M. L., Gibby, R., Milton, O., and Pottharst, K. "The Effect of Varied Experimental 'Sets' Upon Rorschach Test Performance," J. Projective Tech., 1950. 14.

13. Jonietz, A. K. "A Study of the Phenomenological Changes in Perception after Psychotherapy as Exhibited in the Content of Rorschach Percepts," Unpublished Ph.D. dissertation, University of Chicago, 1950.

14. Klopfer, B., and Kelley, D. McG. The Rorschach Technique. Yonkers-on-Hudson, New York: World Book Company, 1946.

15. Koren, L., Goertzel, V., and Evans, M. "The Psychodynamics of Failure in Therapy," Amer. J. Psychiat., 1951, 108, 37-41.

16. Lord, Edith. "Experimentally induced Variations in Rorschach Performance," Psychol. Monographs., No. 316, Washington D. C.: Amer. Psychol. Assoc., 64, 1950.

17. Miale, F. R., and Harrower-Erickson, M.R. "Personality Structure in the Psychoneurosis," Rorsch. Res. Exch., 1940, 4, 71-74.

18. Miller, D. R. "Prediction of Behavior by Means of the Rorschach Test. Dimensions of the Interpersonal Situation," Unpublished paper read at Amer. Psychol. Assn., 1951.

19. Milton, E. O., "The Influence of Varied Experimental Sets upon Certain Rorschach Variables: II. Stability of the Human Movement Variable," Unpublished Ph.D. dissertation, University of Michigan, 1949.

20. Monroe, Ruth. "The Inspection Technique: A Method of Rapid Evaluation of the Rorschach Protocol," Rorsch. Res. Exch., 1944, 8. 46-70.

21. Piotrowski, Z. "A Simple Experimental Device for the Prediction of Outcome of Insulin Treatment in Schizophrenia," Psychiat. Quart., 1940, 14, 267-273.

22. Pottharst, K. "The Influence of Varied Experimental Sets Upon Certain Rorschach Variables: III. The Influence of a Complex Set," Unpublished Ph.D. Dissertation, University of Michigan, 1950.

23. Rorschach, H. Psychodiagnostics, (2nd Ed.) Verlg Hans Huber, Berne, Switzerland. Distributed by Grune and Stratton, New York, 1942.

24. Schrier, Harvey. "The Significance of Identification in Short-term Therapy," Unpublished Ph.D. dissertation, University of Michigan, 1951.

25. Stotsky, B. A. "Factors in Remission of Schizophrenics: A comparative study of Personality and Intellectual Variables Among Schizophrenics," Unpublished Ph.D. dissertation, University of Michigan, 1951