



eCOMMONS

Loyola University Chicago
Loyola eCommons

Dissertations

Theses and Dissertations

1970

Changes in Motivation During Client-Centered Therapy as Measured by the Story Sequence Analysis of the Tat

Joseph Augustine Ramirez
Loyola University Chicago

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

 Part of the [Psychology Commons](#)

Recommended Citation

Ramirez, Joseph Augustine, "Changes in Motivation During Client-Centered Therapy as Measured by the Story Sequence Analysis of the Tat" (1970). *Dissertations*. 1049.
https://ecommons.luc.edu/luc_diss/1049

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



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 3.0 License](#).
Copyright © Joseph Augustine Ramirez

Loyola University

CHANGES IN MOTIVATION DURING CLIENT-CENTERED THERAPY
AS MEASURED BY THE STORY SEQUENCE ANALYSIS OF THE TAT

A DISSERTATION

Submitted to the Faculty of the Psychology

Department of Loyola University

in Partial Fulfillment of the

Requirements for the

Degree of

Doctor of Philosophy

by

J. Augustine Ramirez

Loyola University

Chicago, Illinois

February, 1970

JOSEPH AUGUSTINE RAMIREZ

Biographical Sketch

Was born in Guadalajara, Jalisco, Mexico, on January 13, 1912.

Became American Citizen in March, 1965.

Education

<u>Institution and Location</u>	<u>Degree</u>	<u>Year Conferred</u>
Scotus College, Hebbbronville, Texas	B.A.	1934
Catholic University of America, Wash., D.C.	M.A.	1938
Catholic University of America, Wash., D.C.	Ph.D.	1955

Dissertation: "Unconscious Drives and Human Freedom in
Freud's Psychoanalysis."

University of Chicago, Department of Psychology M.A. 1964

Thesis: "Ethical Development during Client-Centered Therapy."

Professional Experience

Instructor of Philosophy, Scotus College, 1938 to 1945.

Director and Principal of "Instituto Durango," Durango, Mexico,
1945 to 1951.

Post-Doctoral Privileges at the University of Chicago, Chicago,
Illinois, June, 1956 to August, 1958.

Honorary Staff Member of the Counseling Center, University of Chicago,
Chicago, Illinois, September, 1958 to June, 1961.

Staff Member of the Counseling Service of Catholic Charities, Chicago,
Illinois from January, 1959.

Registered Psychologist in Illinois, March, 1965.

Assistant Director of Psychodrama, Moreno Institute, New York, 1965.

Member of the American Catholic Psychological Association since 1964.

Member of the American Society of Group Psychotherapy and Psychodrama since 1965.

Experience in individual, family and group therapy over 10,000 hours.

ACKNOWLEDGMENTS

The author wishes to extend to Dr. Magda B. Arnold, Originator of the Story Sequence Analysis and Advisor of this dissertation, the expression of his admiration for her work and of his gratitude for her assistance in carrying out this study.

A sincere debt of gratitude is also due to Dr. Ronald Walker, Dean of the Psychology Department at Loyola University, and to Dr. Ann Heilman, Readers of this dissertation, for their valuable suggestions to improve the preliminary report and the final draft of this research.

Special mention must be made of Dr. James Mead, S.J. and of Dr. Patricia Dore who contributed generously of their time to analyze and score the TAT records used for the reliability study.

To Mr. Richard Jenney and the Staff of the Counseling and Psychotherapy Research Center of the University of Chicago, the investigator is greatly obliged for allowing him to use the research material and facilities at their disposal.

To Monsignor Thomas J. Holbrook, Administrator of the Social Services of Catholic Charities of the Archdiocese of Chicago, and to Miss Vera Dillon, Supervisor of the Counseling Service, who made this research possible by granting the author a sabbatical year and limiting and rearranging his counseling schedule, the author is greatly indebted for their understanding and cooperation.

To Mrs. Roseann Cyrier and to Mrs. Bonnie Smolik, my warm appreciation for their help over many English hurdles and for typing this manuscript.

TABLE OF CONTENTS

CHAPTER	PAGE
I. INTRODUCTION	1
II. REVIEW OF RELATED LITERATURE	8
III. PROCEDURE.	23
IV. RESULTS.	38
V. DISCUSSION OF RESULTS.	68
VI. SUMMARY AND CONCLUSIONS.	83
BIBLIOGRAPHY	88

LIST OF TABLES

Table 1 - Personal data and MI scores of the experimental Ss.	41
Table 2 - Personal data and MI scores of the control Ss	42
Table 3 - Wilcoxon Matched-Pairs Signed-Ranks Tests	44
Table 4 - Mann-Whitney U Tests.	53
Table 5 - Rank Order Correlations between the MI scores and Self-Ideal Self r's of the experimental and control Ss	59
Table 6 - Rank Order Correlations between the MI scores and Self-Ideal Self r's of success and failure experimental Ss.	61
Table 7 - Rank Order Correlations between the MI scores and E-M ratings of the experimental Ss	63
Table 8 - Rank Order Correlations between the MI scores and E-M ratings of success experimental Ss	64
Table 9 - Rank Order Correlations between the MI scores and E-M ratings of failure experimental Ss	65

LIST OF FIGURES

Figure 1.	Median MI scores of the experimental and control subjects	47
Figure 2.	Median MI scores of success and failure experimental Ss.	51
Figure 3.	Median MI scores of success experimental and control Ss	55
Figure 4.	Median MI scores of failure experimental and control Ss	56
Figure 5.	Median MI scores female and male experimental Ss.	71
Figure 6.	Median MI scores of female experimental and female control Ss.	74
Figure 7.	Median MI scores of female success experimental and female control Ss.	75
Figure 8.	Median MI scores of female success and male success experimental Ss.	77
Figure 9.	Median MI scores of male success and failure experimental Ss.	78

CHAPTER I

INTRODUCTION

Studies on the changes effected by therapy have yielded inconclusive results. The cause of the equivocal evidence may be attributed partly to the difference of goals which each therapist sets for himself according to his preferred theory of therapy and personality, and partly to the possible compounding of the theory of therapy held by the investigator with the instruments devised ad hoc to measure its effects on the patients. Some investigators have been very careful to make explicit their goals of therapy, the variables to be measured and the devices and statistics they will use to test their hypotheses but unfortunately have neglected the all important independent variable - by lumping together counseling, social work and psychotherapy as practiced by people of different training and orientation who agreed only on very broad general goals. (Volsky et al ... 1965). Other investigators have limited the number of so-called therapeutic sessions to three or four; this limited contact with the therapist can hardly be called therapy and cannot be expected to effect any enduring personality change in the counselee, client or patient. (Morton, R., 1955; Williams, J., 1962). Some studies have tested as outcomes of therapy variables of doubtful validity like temporal perspective (Ricks, D. et al., 1964), affective complexity (Henry

& Shlien, 1958), and verbal productivity (Ullmann & McFarland, 1957). Still others have used criteria of low reliability and have failed to use adequate controls.

Often the outcome of therapy is measured by means of patients' self-reports. These may be influenced by their desire to obtain help, to please or attack the therapist, or by their personal need to convince themselves that therapy was worth their time, money and endeavors. Often, the items of Q-sorts and questionnaires are taken from statements of patients who have gone through the same type of therapy as the one under study. This practice may introduce a bias in favor of its assumptions, hypotheses and results.

While subjective tests thus have a built-in bias, more objective tests like the TAT have other disadvantages. If the test is interpreted subjectively and intuitively, there cannot be adequate interjudge reliability, and the clinical diagnosis derived from the study may be only a personal or theoretical projection of the experimenter. On the other hand, the TAT may be scored objectively by measuring the reaction time or counting the number of words of a definite classification, or by using other indices of content analysis which can be scored easily with high reliability but which are usually irrelevant to therapy. The scores thus obtained, objective and reliable as they may be, have no clear relation to therapeutic success or failure as viewed by the therapist.

To eliminate the variability introduced by different therapists

with different training and orientation, this study will be limited to client-centered therapy as practiced by trained therapists of the Counseling Center of the University of Chicago. Only subjects with at least six therapy sessions will be taken into the experimental group.

To prevent the influence of ideas and assumptions of client-centered therapy from entering in the evaluation of the outcome of client-centered therapy, the change in motivation effected during therapy will be measured by a more objective test like the TAT analyzed and scored by Arnold's Story Sequence Analysis (SSA). This method was developed independently of client-centered therapy and has been validated with high and low achievers in school, business and management. It has been found that a Motivation Index above 100 is characteristic of high achievers. The Story Sequence Analysis will be applied for the first time to measure the outcome of therapy. It is hypothesized that successful therapy cases will have a M.I. score above 100 after therapy is completed.

Since the goal of therapy in general is to effect some enduring changes in the personality of the patient, the ideal criterion of the effectiveness of therapy should naturally be sought in the overt behavior of the subject after therapy. But here again we would meet the lack of agreement among personality and social psychology theorists as to what constitutes the desirable and undesirable aspects of human behavior. In this study the basic criterion of success or failure

in therapy will be the rating by the therapists at the end of therapy on a nine-point scale with 1 indicating complete failure and 9 marked success. It is based exclusively on the therapists' view of their clients during therapy; they have no knowledge of the scores their clients obtained in other psychological tests. The test administrators were different from the therapists. (Seeman, 1954). Though we do not have as yet a validity study of therapists' judgments of success and failure, the ratings done by the therapists have proven to be the most reliable, the most constant and the most uniform measure of outcome used in studies of client-centered therapy (Cartwright et al., 1963; Fiske et al., 1964).

The dependent variable (motivation) has been disparately defined by the contributors to the Nebraska Symposia on motivation. In this study, motivation is defined as the organized pattern of the patient's motives. Motives are his habitual action tendencies, that is, the way the patient ordinarily faces a particular problem and the way he habitually solves it. Motives are an action set that determines his appraisal of the problems and situations he faces and the choice of his preferred solution under the influence of his needs, his imagination, his ego reorganization, his past experiences and his learned reactions to approach and avoidance. Our dependent variable is unaffected by the many controversies regarding the origins of human motivation. The method of Story Sequence Analysis displays the motivational attitudes at the moment of the TAT test whether they are

conscious or unconscious, biological or learned, approach or avoidance. Operationally, motivation is defined as action tendencies of the patient as expressed in the imports of the TAT stories in four areas of concern: I. Achievement, success or failure; II. Right or wrong from the rational and ethical viewpoint; III. Human Relations; and IV. Reaction to adversity. (For description of the Ss, and the scoring system of the Story Sequence Analysis, cf. Chapter III: Procedure).

The concrete working hypotheses of the present study are:

1. Experimental Ss who received client-centered therapy for more than six sessions will show a significant improvement in their motivational attitudes at the end of therapy and follow-up, i.e., will have a significantly higher M.I. score, than at the beginning of therapy.
2. Control Ss who took the TAT at approximately equal intervals as the Experimental Ss but who did not receive therapy, will not show a significant improvement in their motivational attitudes, i.e., will not have a significantly higher M.I. score on their second and third testing corresponding to post-therapy and follow-up, than in their first test corresponding to the beginning of therapy of the experimental Ss.

Since not every patient who undergoes therapy is successful in his efforts to change his personality or to solve his problems, the beneficent effects of therapy are often clouded by the negative results

of the failure cases. To clear up this confusion, the experimental group will be divided into success and failure subgroups by the rating of their therapists. Patients who received a rating of 6 or more by their therapists at the end of therapy will form the success subgroup; patients who received a rating of 5 or less at the end of therapy will constitute the failure subgroup. These subgroups will be compared with each other and with the control Ss at the beginning, end and follow-up period of therapy and/or testing.

3. Patients rated as success cases by their therapists will show a significant positive change in motivation, i.e., will have a significantly higher M.I. score, at the end and follow-up than at the beginning of therapy.

4. Patients rated as failure cases by their therapists will show no significant positive change in motivation, i.e., will not have a significantly higher M.I. score, at the end and follow-up than at the beginning of therapy.

5. Finally, it is our hypothesis that success experimental Ss will be better motivated than failure cases, i.e., will have a higher M.I. score, even at the beginning of therapy.

It is hoped that by comparing the M.I. of the success and failure cases at the beginning of therapy it is possible to establish a cut-off point which, together with indications derived from other psychological tests, could be used to predict the outcome of therapy for individual patients as success or failure. If this quest is successful, it

would help relieve the strained facilities of counseling agencies by diagnosing applicants who are poor therapy risks.

In almost every study of the outcome of client-centered therapy, the investigator has been asking himself these questions: Is the change effected by successful therapy a modification of the entire personality of the client or does it change only his self-concept and the various aspects of his personality touched and brought out by the different psychological tests used? How do these various tests stand in relation with each other? (Cartwright, D. et al., 1963). A tentative and partial answer to these intriguing questions will be given by comparing the M.I. scores of the experimental and control subjects in this study with the scores they obtained on other psychological tests. The tests to be compared with Arnold's Story Sequence Analysis are Butler-Haigh Self-Ideal Self Q-Sort, Willoughby Emotional Maturity Test scored by the patients themselves and two observers, the Personal Integration Rating of the therapists and the TAT analyzed and scored clinically by an independent diagnostician.

CHAPTER II

REVIEW OF RELATED LITERATURE

It was the firm belief of the creators of the Thematic Apperception Test that the stories given to the TAT cards would reveal the repressed dynamic needs and presses acting on the individual at the time of the test, thus giving the therapist a quicker insight into the nature and causes of his patient's problems. They wrote:

If the analyst were cognizant at the very start of the fundamental fantasy constructions of his patient he should be in a better position to apperceive and to interpret the dynamic relations of what, in the beginning of an analysis, is ordinarily fragmentary and obscure. He might also, at a later stage, have a better idea of what might be considered irrelevant as well as what important latent trends had yet to be disclosed. (Morgan & Murray, 1935, p. 62)

They were equally confident that the Thematic Apperception Test with "a ten page autobiography, an hour of relaxed reminiscing about childhood experiences... and an hour of direct questioning" could provide people in distress and in need of guidance with a short economic analysis that would help them along in their problems and progressive efforts (Ibid.)

Following Murray's suggestions some clinicians have used the TAT as a valuable auxiliary technique during therapy. Bellak (1950) recommends the use of the TAT in therapy as a means of "communication, interpretation, insight and working through", especially when the patient has difficulty in focusing on the coming to grips with his real problems, and instead keeps talking about irrelevant or superficial

problems. Deabler (1949) would ask the client to take the TAT and let him talk freely about the stories with the permissiveness and reflection of feelings characteristic of client-centered therapy. Bettelheim (1947) demonstrated the use of the TAT as a teaching and therapeutic device. He would ask the students to take the TAT and then help them analyze it by writing down every personal reaction and association that came to their minds as they read each one of their TAT stories. Arnold suggested as far back as 1949 a short clinical analysis of the TAT as a quick and reliable means of diagnosing the main problems of patients. She would count the number of stories written along five social categories, and point by their sequence and consistency the relevance to the pressing problems of the patients (Arnold 1949). This seems to be the beginning of the method of Story Sequence Analysis (SSA), which will be explained in a subsequent chapter. Some of Arnold's students use also the SSA in connection with therapy. They ask their patients to take the TAT at the beginning of therapy or when their relationship has been well established. The patient writes five-minute stories to the TAT and the therapist writes the imports of the patient's stories in the first person. He then hands a copy of the stories and the imports to the patient and asks him to comment on them. They report that the SSA gives the patient and the therapist new insights into the patient's problems and helps him talk about them in a more open and personal way.

Some clinicians have gone a step farther and have used the TAT as a form of short therapy after Murray's suggestion. Morton (1955) used the TAT as a form of brief therapy and measured its results by Mooney Check List and Rotter's Sentence Completion Test. His subjects were 40 undergraduates who had applied for therapy. He matched them on a maladjustment score derived from Mooney's Check List and Rotter's Sentence Completion Test before assigning them at random to either the experimental or control group. The experimental group wrote stories to a TAT in their first interview and took them home to read them over and think over their problems. In the next two sessions they discussed with the therapist the problems revealed by the TAT. Finally, all Ss, experimental and control, took again Mooney's and Rotter's tests. Morton found a significant decrease in the maladjustment score in the experimental group but not in the control. He claims that the TAT administered and discussed this way is an effective form of brief therapy as shown by the difference in the maladjustment score between the two groups at the end of therapy. There is no doubt that taking the TAT, mulling over the stories, and discussing briefly with a therapist the problems uncovered by the test, has an impact on the subject. However, two sessions with a therapist would hardly be called therapy in the strict clinical sense. The validity of the measures used is also questionable (Cross, 1964).

Our interest in the TAT in this study is not in its validity as a therapy substitute nor its efficiency as a auxiliary technique in

therapy, but rather as a reliable method to evaluate the outcome of therapy as understood and practiced by professional therapists. Brief reference was made to related uses of the TAT in order to understand some of the results to be reported later in the present study.

Meyer and Tolman (1955) hypothesized that the TAT protocols at the beginning of therapy could predict the content and affective reactions of patients in their therapy sessions. They gave at the beginning of therapy a 6-card TAT to 50 patients, 33 males and 17 females. The investigators analyzed both the TAT stories and ten of the recorded therapy sessions, ordinarily the first ten, by means of check lists of words descriptive of parental figures and of the qualities attributed to them. A different check list was used for the descriptions and attitudes of father and mother or of father-and mother-substitutes. The results contradicted their hypothesis. The content of the recorded interviews did not agree with their scoring of the TAT by means of check lists. The interjudge reliability was .84 for the therapy sessions but only .54 for the TAT stories. The rather low reliability for the TAT stories in opposition to the reliability reached for the actual therapy sessions stands out and points clearly to the subjectivism introduced into the TAT protocols when analyzed in terms of parental substitutes so obviously subject to the vagaries of each interpreter. The main defect of this study, however, lies in the unwarranted assumption of taking the hero and the

central figures of the TAT stories as a source of autobiographical information.

Ricks, Umbarger and Mack (1967) argued that delinquency originates in the impulsiveness of the delinquent who disregards the history of the problem and the consequences of his acts. They defined delinquency as the narrowing of the time perspective in the mind of the delinquent, and hypothesized that it would show and could be measured in the TAT stories. They used for this purpose a 15-card TAT designed to tap changes in self-image, control of aggression and attitudes toward authority. The investigator had ten experimental and ten control delinquent boys write stories to this special TAT before and after vocational guidance and therapy given the experimental Ss for eleven months. The controls took the TAT but had no special treatment. The 600 TAT stories of the experimental and control groups were rated by two scorers on an 11-point time scale with concrete indications of time from less than an hour to a number of days, weeks, years and the whole life span. The raters were asked to figure out and score the time implied in the TAT story before and after the situation suggested in the card and described by the storyteller. The interjudge reliability was fairly high, .80 for prospective and .81 for retrospective time. The investigators report that after therapy the treated boys increased their prospective TAT time in the control of aggression and in the self-image, but not in the attitudes toward authority. They also report increase of the retrospective TAT

in the control of aggression but not in self-image nor in the attitudes toward authority. Narrowing of time perspective in the mind of the delinquent boy may be a plausible factor of their delinquent behavior, but it is not clear how the time increase in the TAT stories, as measured by the investigators, should be referred to therapy. For one thing, the investigators bunched together vocational guidance and therapy without indicating what kind of therapy was given and whether it was successful or unsuccessful according to the therapists. Since they found a high correlation between the time perspective and the scores of the subjects in tasks which are definitely a function of intellectual ability like achievement in mathematics and ability to solve arithmetic problems, it is probable that the variable they were measuring as dependent variable was a product of the Ss intellectual ability and vocational guidance rather than an outcome of therapy. Another flaw in the study which casts doubt on the results they report, is that the investigators mention the inter-judge reliability in the measurement of prospective and retrospective time but not in the categories of control of aggression, self-image and attitudes towards authority used to report the results. Since the reader is not informed of the reliability and validity of these categories in relation to the time perspective that is being investigated, he has some basis to doubt the results as reported.

Henry and Shlien (1958) investigated the outcome of limited therapy (20 sessions) and unlimited therapy (voluntary termination). They used as dependent variable "affective complexity" which they defined as the

ability to recognize opposite feelings in oneself and in others. They assumed that patients who attribute only one feeling to the same person are repressing the opposite feeling. From this a prior assumption they concluded that successful therapy should remove this repression and enable the patient to express opposite and contradictory feelings about themselves and others in a significantly higher degree than at the beginning of therapy, even in their TAT stories. They gave a 5-card TAT to patients before and after client-centered therapy. To measure the variable of affective complexity in the TAT stories, one scorer rated each TAT protocol on four different scales or indices from which a combined composite score of affective complexity was derived. The intrajudge reliability was tested by rescored 25% of the records six months later. The reliabilities obtained ranged from 68% to 94% in the four indices. Contrary to their hypothesis, they found no significant increase in affective complexity from pre-therapy to post-therapy in either group. From pre-therapy to follow-up the unlimited therapy group showed significant increase in the affective complexity score but the limited therapy group showed an equally significant decrease. The investigators are at pains to explain their rather negative results. There are several methodological flaws which may have contributed to their failure. There was only one scorer, and his intrajudge reliability varied from index to index. Moreover, the four indices used as measures had a different range, from 0 to 3, 1 to 3 and 1 to 5, thus contributing different weights to the composite score. Finally, the

rater gave one global score to each record in each one of the scales thus adding the vitiating factor of subjectivity to the obvious lack of precision in the concept of affective complexity and instability of instruments used to measure it.

Carl Rogers and his collaborators (1954) produced a monumental work for the measurement and evaluation of the outcome of client-centered therapy. Employing different measures, most of the investigators found significant positive changes in the experimental subjects from pre-therapy to post-therapy and from pre-therapy to follow-up, 6 to 12 months after the termination of therapy. These positive changes were not observed in the control Ss who took the same battery of tests at matched time intervals but who did not receive therapy. However, some of the tests are based on the subject's view of himself, which is not necessarily valid. Also, it is likely that a test such as the Self-Ideal Self Q-Sort, developed from statements of patients in client-centered therapy, might introduce a bias in favor of the therapy which it is supposed to evaluate.

Realizing that a projective test like the TAT is preferable to self-ratings and self-report because it cannot be manipulated by the patient in the same way, Dymond (1954) used a blind analysis of the regular 20-card TAT to evaluate the outcome of client-centered therapy. She obtained 72 TAT records from 25 experimental Ss from the beginning, end, and follow-up, and 20 records from 10 control Ss who had no therapy but took the TAT at matched time intervals. The investigator scored the

TAT records on a 7-point scale of mental adjustment giving one single global score to each record. Her intrajudge reliability was .94. She found a significant improvement in the scores of the experimental group from pre-therapy to post-therapy and from pre-therapy to follow-up, which was not shown by the control subjects. However, contrary to her expectations, the experimental group had a significantly lower score than the controls at the beginning of therapy, and at the end of therapy they did not differ significantly from the controls. These are important results because they show that the TAT can be used with consistency for the measurement of the outcome of therapy. Her study was weakened, however, by the subjectivity of the scoring system used. Since there was only one rater, and she was a client-centered therapist and members of the staff in the research project, it is possible that the TAT protocols were scored on the basis of subjective criteria derived from her counseling experience and her collaboration in the whole program of research. The element of subjectivity is enhanced by the global score given to each protocol, which could hardly be matched if other scorers had been used.

In a more recent study, Cartwright and Vogel (1960) made a considerable improvement on the use of controls in psychotherapy research. By an ingenious device, they took each subject as his own control and tested the effects of therapy on a period of time equal to the waiting period of each client. Twenty-two subjects took the Self-Ideal Self Q-Sort and the TAT when applying for therapy, again at the

beginning of therapy after a waiting period that varied from 3 to 24 weeks according to the availability of therapists, in therapy at a point equal to their waiting period and finally at the end of therapy. All the therapists followed the client-centered orientation but varied in experience from 3 to 60 cases. The main hypothesis of these investigators was that their subjects would show more change and improvement during therapy than during an equal matched waiting period, and that their score on both tests would be significantly higher after therapy than at the beginning of therapy. The TAT was scored by only one investigator on a rating scale equivalent to Dymond's former 7-point mental adjustment scale. In general, their results were negative. The Ss did not show any significant improvement in either of the tests used at the in-therapy period equal to their waiting time nor at the end of therapy. Only when the investigators divided the patients into a long- and short-wait subgroup (2-7 weeks vs. 6-24 weeks) and the therapists into an experienced and inexperienced groups according to the number of cases they had treated prior to this study (2-5 vs 6-60) did they find some significant but also some disconcerting and contradictory results. On the TAT test, only the patients under experienced therapists showed significant improvement in their in-therapy score and at the end of therapy. No improvement in the TAT score was shown by patients working with inexperienced therapists at the in-therapy test nor at the end of therapy. On the other hand, patients working with inexperienced therapists evidenced significant improvement in their Q-Sort at the in-therapy

period but not at the end of therapy, while patients working with experienced therapists showed significant change in their Q-Sort at the end of therapy but not at the in-therapy testing period. These results are significant on two counts: first, they show once more that the experience of the therapist is an important factor determining the successful outcome of therapy, and second, they show that the TAT is less amenable to subjective influences and expectations than the Q-Sort. It was only patients with inexperienced therapists who showed significant change in their Q-Sort at the in-therapy period. That it was artificially determined by factors other than therapy is evidenced by the fact that this improvement did not hold at the end of therapy. The investigators honestly report an unexpected result which baffles them and throws some shadow on the positive results obtained in the TAT score by patients working with experienced therapists. When the investigators analyzed the change during the waiting period by the short-and long-wait subgroups, they found that patients who had been waiting longer showed significant improvement in their TAT before therapy, while patients with a short-waiting period had a significant deterioration in their TAT scores before therapy. The Q-Sort did not show any difference between these subgroups during their waiting period. The improvement and deterioration shown in the TAT scores during the waiting period again demonstrates that the TAT can have a significant impact on the testee without therapy. Why improvement should appear in people in longer waiting period is not clear and has to be studied by

an investigation designed for this purpose. Though this study solves the thorny problem of controls in therapy research by an ingenious device, it is vitiated again by the subjectivity of the scoring system which is perhaps at the root of the negative and contradictory results obtained. There is also a noticeable silence on the intrajudge reliability of the one scorer used to rate the TAT protocols.

In an attempt to avoid confusing the assumptions and expectations of client-centered therapy with the rating of its outcome, Grummon and John (1954) devised twenty-five 7-point scales based on psychoanalytic theory to score the TATs of experimental Ss at the beginning, end, and follow-up of therapy, and the TATs of control who had no therapy but took the TAT at matched time intervals. They found a significant improvement in the scores of experimental Ss from pre-therapy to post-therapy, and from pre-therapy to follow-up on most of the scales, particularly of Ss who had high initial scores at pre-therapy. There was no significant change in the scores of the controls. Though this study was successful in showing that the outcome of client-centered therapy could be evaluated by tests based on psychoanalytic theory, thus avoiding an important source of bias, it was not successful in avoiding subjective bias as shown by the global intuitive score given to each TAT record on the 25 scales, and furthermore by the fact that two scorers did not reach adequate interjudge reliability.

Finally, Barrington (1967) used the TAT to evaluate the outcome of client-centered therapy with patients of a state mental hospital. He

scored blindly 66 protocols from an 8-card TAT given at the beginning and end of therapy. Twenty-four records were from twelve experimental Ss diagnosed as schizophrenics; twenty were from ten control Ss from the same hospital matched with the experimentals in sex, age and length of residence in the hospital, and twenty-two were from non-hospitalized normal people who volunteered in the research program and evidenced no need for psychological or psychiatric treatment. The investigator devised eleven scales which he thought would be related to the main constructs and assumptions of client-centered therapy like openness to experience, decrease of emotional distance, sensitivity and use of own feelings. He scored the 66 protocols given one single global score for each one of the variables of the eleven scales. Fifteen randomly selected protocols were rescored on six scales one year later. The obtained intrajudge reliability on the six scales ranged from .78 to .96. The scores of the experimental and control subjects were compared singly on each one of the eleven scales from pre-therapy to post-therapy and by one composite score derived by the combination of the scores of the eleven scales. The experimental Ss showed significant improvement from pre-therapy to post-therapy on four of the scales while the hospitalized controls showed significant improvement in only one of the scales; the nonhospitalized normal controls showed a significant decrease in one of the scales. By combining the scores of the eleven scales, the experimental Ss showed some improvement from pre-therapy to post-therapy, but it is not statistically significant nor statistically

different from the improvement shown by the hospitalized controls. However, by adding the number of categories for each individual where improvement was noted and subtracting the number where negative change was shown, each subject came off with a positive, neutral or negative score. Ten of the experimental subjects had positive scores and only one had a negative score. From the hospitalized group, six had positive and four had negative scores; the nonhospitalized group had three with positive and seven with negative scores. The investigator shows that this distribution of the experimental group is significant at the .05 level. Nonetheless, the overall result of this study is rather negative and would indicate that the categories chosen and the scales devised by the investigator to measure them refer to different areas of experience and change which tend to cancel each other. Their overall correlation with the therapists' rating of the outcome of therapy was $-.11$, which is slightly negative and not significant but indicates that the categories chosen and the measures devised ad hoc lack construct validity. Only one of them -- decrease in emotional distance -- correlates significantly with the therapists' rating of success. This study is further weakened by the subjective method of scoring the TAT by one single global score on variables or categories whose reliability and validity have to be proven.

In the present investigation we propose to use as independent variable an accepted form of therapy as practiced by trained therapists of the University of Chicago, thus avoiding the questionable use of the

TAT as a form of brief therapy and the confusion of therapy with lectures on vocational guidance. To be sure that the Ss have been exposed to the independent variable only Ss with at least six therapy sessions will form the experimental group. The dependent variable - motivation - defined before as attitudinal action tendencies, is broad enough to account for improvement in every area of human activity, and should be more valid as a measure of the outcome of successful therapy than the limited and questionable variables used in some of the studies we have reviewed. Great care will be exercised to insure adequate objectivity and reliability to the TAT scores by following Arnold's objective method of Story Sequence Analysis and by scoring each story with the help of appropriate imports and the detailed instructions and concrete anchors of the scoring system (Arnold, 1962).

CHAPTER III

PROCEDURE

A. Experimental Subjects

The experimental subjects are thirty out-patients at the Counseling Center of the University of Chicago. They have never been hospitalized in psychiatric institutions. They applied for therapy and agreed to participate in this study as research in personality, and to take the TAT and a battery of other psychological tests at the beginning, end, and follow-up periods of therapy. Nineteen are male, and eleven are female; their age ranges from 20 to 40 years with a median of 26.8 years. All have middle class status (range from lower to upper middle). They are above average in intelligence and education with college and post-graduate background; sixteen are students at the University of Chicago and fourteen are in different professions and occupations in the community. An untrained observer could perhaps observe in them some idiosyncratic but not serious emotional reactions. A diagnostically oriented clinician would classify most of them as psychoneurotics, some as borderline psychotics and a small group as normal with slight emotional disturbances due to environmental problems and tensions.

B. Control Subjects

Thirty control subjects were matched with the experimental Ss in age, sex, education status as student or non-students and socioeconomic status. Their ages ranged from 19 to 44 years with a median

of 27.0 years. Fifteen were male and fifteen female; fifteen students and fifteen non-students. The control Ss agreed to participate in this study, and to take the TAT and the battery of other psychological tests at approximately the same time interval as the experimental subjects, but they were not seeking therapy nor were they receiving psychotherapy elsewhere.

C. Therapists

After taking the first TAT and other psychological tests, the experimental subjects were assigned to twelve therapists who had been trained in and purportedly were following the client-centered orientation. Five of the therapists were considered less experienced with less than four years of practice; seven therapists had from 4 to 21 years of experience with a median of 5 years. At the end of therapy all the therapists rated the outcome of therapy on a 9-point scale: 1 indicating complete failure and 9 marked success. They had no knowledge of the results of any of the psychological tests, and rated their clients exclusively on their observation and the interaction with them during therapy. It can be assumed, however, that they were influenced in their ratings by the theory of client-centered therapy which considers therapy as a corrective emotional experience and a personal exploration of one's problems rather than intellectual guidance, cognitive discussion and evaluation of situational and environmental problems. Seven cases were rerated five months later to check the reliability of the therapists' ratings. The mean correlation between

the ratings was .81, indicating a rather high degree of consistency and stability in their ratings in spite of such a long interval of time (Seeman, 1954). Studying another sample of cases Cartwright (1955) asked eight counselors to rerate 15 clients. The mean length of time between the first and second ratings was 14.2 months, a span over which considerable forgetting might have been expected. The rate-rerate reliability was .86. This result, with that of Seeman, may be regarded as evidence that the nine-point counselor rating scale of success in therapy is a reliable instrument. The therapist's rating of the outcome of therapy will be the criterion used in this study to discriminate between the success and failure cases among the experimental group.

D. The Research Material

The regular 20-card Thematic Apperception Test with Murray's original instructions and with the BM and GF cards varying according to the sex of the subject (Murray, 1948) was given at the beginning, end and follow-up periods of therapy or equivalent time intervals for the controls. The test was administered individually in two sessions for the first and second 10-card sets with a day or two of separation. The stories were electrically recorded and later transcribed by the clerical staff of the Counseling Center of the University of Chicago. The test administrators were not therapists, but they were well trained and experienced in the administration of the test and were able to establish warm relations with the testee while maintaining a psychological and professional distance. Though all the subjects, experi-

mental and control, took the first TAT at the beginning of therapy, some failed to meet their second and third testing appointments especially among the control group; four tapes were not transcribed because, due to mechanical defects of the machine, they were completely inaudible or spotty.

After the blind analysis was completed by the investigator, some of the records could not be identified as belonging to either the experimental or control group. (The reasons for this will be explained later). Therefore the actual number of TATs that were used for the study was reduced to 156, or 87 TATs from the experimental and 69 from the control subjects. This will be indicated in the analysis of data and in the report of the results of the study.

The mean interval of time between the first and the second TAT at the end of therapy was 170 days for the experimental group and 178 days for the controls; the mean interval of time between the second and the third TAT as follow-up was 203 days for the experimental and 219 days for the control group.

E. Method of Analysis and Scoring System

The TAT records were analyzed and scored by Arnold's method of Story Sequence Analysis (SSA) along four categories, with the weights of the scoring system explained and illustrated in her book (Arnold, 1962). The beginnings of the method go back to 1949 when Dr. Arnold was Director of Research and Training in the Psychological Services of the Department of Veteran Affairs in Canada. It has been refined and made more precise by the empirical studies done in recent years in collaboration with her

students.

1. Rationale of the Method

The method is based on Arnold's idea that the TAT stories are a product of the imagination guided by the estimative dispositions and emotional action tendencies of the individual. They are spun by the imagination from memory images of past perceptions and personal experiences but they are not memories. The TAT stories betray the storyteller's way of meeting a particular problem and his personal convictions about human actions by the plots and outcomes he puts into his stories.

Following Casson, Arnold defines human personality as "the patterned totality of human powers, activities and habits, uniquely organized by the person in the active pursuit of his self-ideal, and revealed in his behavior". (Arnold and Casson, 1954, p. 219). The TAT stories are certainly a sample of human behavior, and it is reasonable to assume that they will bear the marks of the organization and consistency of all the powers, dispositions and action tendencies of an individual at a particular stage of his life and development. "The stories that a man tells, like the dreams he has, illustrate problems that occupy him, solutions he is working out and convictions he has achieved." (Arnold, 1962, p. 25).

Essential to the rationale of the Story Sequence Analysis is Arnold's concept of the import of the TAT story. Challenged by the stimulus card, the storyteller sketches a plot that usually represents one of his problems; the outcome reflects a tentative solution that

seems appropriate to him. An import is not a summary of the story; rather, it presents the moral of the story in the subjective circumstances of the storyteller. It reveals the trend of thought of the storyteller and the way he evaluates human actions.

The story import will show how the storyteller thinks people usually act and how he feels they should act; what actions he thinks right and which wrong; what will lead to success, in his opinion, and what to failure; what can be done when danger threatens, and what are the things to strive for. In short, the story imports, taken in sequence, give us a connected statement of the storyteller's principles of action, his motivational patterns. Obviously, this pattern should make it possible for us to gauge how he would react to a situation. (Arnold, 1961, p. 51).

Arnold enumerates the advantages of using imports to score the TAT stories;

a. Since it is the import that is scored as positive or negative, and the import abstracts the plot and outcome, each picture has the same chance of yielding a positive or negative score.

b. The story sequence analysis provides a new dimension in clinical evaluation. It is almost a self-recording portrait of the storyteller which tells its own story without elaborate and often speculative interpretations.

c. Scoring story imports instead of themes also equalizes the story length. Whether long or short, every story is reduced to an import which is usually contained in one sentence.

d. Finally, this method reveals positive or negative motivation as expressed in vocational adjustment and achievement. This motivation is stable. The positive or negative attitudes revealed in successive

sets of stories to a TAT given within an interval of one year have been shown to remain similar and to obtain similar scores, even though the actual imports were quite different (Arnold, 1962, p.15).

"The import with the sequence in which it is embedded is the backbone of our method of interpretation" (Ibidem, p. 64). The import is objective in the sense that it is abstracted as accurately as possible from the TAT story without adding any subjective or theoretical interpretation. However, it is subject to the unconscious influence of the examiner's focus of evaluation regarding the problem around which the story is spun; the imports as well as the sequence in which they are embedded may vary from examiner to examiner. Some TAT stories are simple and straight-forward, and the storyteller's outlook on life, his goals and the means to attain them, his ideas of right and wrong, his evaluation of human relations, and his reactions to adversity, are crystal clear in the plots and outcome of his stories. There are, however, other stories where the import can be scored in different categories, and the examiner has to make a choice according to the sequence of imports that is most obvious to him.

In the present study it is assumed that the motivational attitudes of the control subjects who took the TAT but had no therapy will remain basically the same from the first to the second as well as from the second to the third TAT, while it is hypothesized that the motivational attitudes of the experimental subjects will change from negative to positive, so that their motivation index will increase as a result of

client-centered therapy.

2. Scoring System

Each import is scored in one of four categories: I. Achievement, success, happiness, active effort or their opposites; II. Right and wrong; III. Human relations, and IV. Reactions to adversity. These categories were empirically derived from imports of high and low achievers at school, from effective and ineffective teachers, from offenders and nonoffenders among Navy recruits, and from efficient and inefficient executives in a government project. Various headings and subheadings with appropriate anchors are given under each category to make the scoring of each import more reliable.

In the present stage of development of the SSA there are four possible scores ranging from very positive to extremely negative: +2, +1, -1, -2. In general, +2 is scored when the import describes overt and positive action: +1 indicates activity that is positive but may not be overt (v.g., planning, positive attitudes) or activity that is not very positive (e.g., there is failure along the way but success is reached eventually). Lack of positive action is indicated by -1 (e.g., success is the result of passive dependence on others; failure comes because nobody helps or advises) while -2 indicates frankly negative, impulsive, or malicious actions or attitudes (e.g., failure leads to desperate action; success comes in spite of refusal to take reasonable advice; wrong succeeds, etc.) (Arnold, 1962, p. 106).

The algebraic sum of all imports, positive and negative, is trans-

formed into a composite score called Motivation Index (M.I.) by the following formula:

$(\frac{\pm n^o}{n^p} 200) + 100$, where n^o is the number of units obtained (derived from the algebraic sum of the scores of the imports) and n^p is the number of units obtainable or possible. The number of possible units for a 20-story sequence is 80, and for an 11-story TAT is 44. Arnold has worked out a table to make this transformation easy. (Arnold, 1962, pp. 46-47).

3. Reliability Method

The reliability of the Story Sequence Analysis was tested by Arnold by the odd-even and split-half methods. She scored 99 TAT records of efficient and inefficient teachers, and 51 TAT records of seventh-grade children using the latest improved scoring system (Arnold, 1962, Appendix). She found a correlation of .86 between the odd and even numbered imports of the teachers' records as well as between the first and second halves of the records. The correlation between the odd and even numbered imports of the children's was .79, and the correlation between the first and second halves was .61. The difference of results with adults and children may be explained by the composition of the two groups. The teachers' group comprised only the extremes of the sample: the high-and low-rated teachers; while the children's group included also the midrange of the sample. Besides, it is easy to conjecture that the motivational attitudes and action tendencies of seventh-grade children are not as well formed and consistent as those of adults, and that they would shift in their motivation from story to story, especially toward the end of the test, as they

explored alternative solutions to the problems presented by the TAT cards.

4. Validity

The scoring system of the SSA has been developed on the basis of studies that distinguished between matched pairs of high and low achievers (Snider, 1954) Brown, 1953, McCandlish, 1958). In these studies, the middle range was excluded. When the imports of stories that were given by high achievers were scored as plus, and the imports of stories given by low achievers as minus, and the imports were arranged in categories, the import of any given story could then be scored as plus or minus and the system could be used to correlate TAT scores with an outside criterion. This preliminary scoring system discriminated reliably between high and low achievers (McCandlish, 1958), successful and unsuccessful teachers (Burkard, 1958), offenders and non-offenders (Petrauskas, 1958). Still later, an attempt was made to distinguish between different intensities of positive or negative motivation by using a score of 1 for strongly negative and a score of 2 for mildly negative imports, while a score of 3 was given for mildly positive and a score of 4 for strongly positive imports. Using this scoring system, Garvin (1960) found a correlation of .85 for men and .83 for women between their TAT scores and their grade point average in college; and Quinn (1962) obtained a correlation of .61 and .59 between his TAT scores and the ratings of possible success of scholastics by superiors and peers. Still later, scores of -1 and -2 were used for negative imports, and +1 and +2 for positive imports. This is the scoring system described in Arnold's manual (1962). With

this scoring system, Arnold (1962) found a correlation of .75 between M.I. and grade point average of fifth grade children, and Dulin (1968) found a correlation of .83 between M.I. derived from TATs given at the beginning of freshman year and the grade point average at the end of the year. As far as school achievement is concerned, the validity of the scoring system either in its preliminary or its later version seems to be well established. Finally, Najjar (1967) designed a study to determine the effectiveness of the SSA in differentiating the motivational patterns of two groups of psychopaths selected on the basis of their MMPI Pd scores and Astin's Self-Esteem Factor Pd scale. He found some significant differences in categories II and III which he used to elucidate the qualitative differences of the two groups. He also tested the interjudge reliability of two raters obtaining a correlation of .74. He acknowledged the limitations of his study which suffered from the lack of appropriate normal controls.

5. SSA and Therapy

In the present study, the investigator has endeavored to go a step farther, and has subjected the Story Sequence Analysis to a more critical test. Other studies have shown the efficacy of the SSA to bring out the habitual motivational attitudes of normal people; this study is testing the adequacy of the SSA as a measure of motivational changes effected by client-centered therapy using the same categories empirically derived from subjects who were not in therapy.

F. Analysis of Data

156 TATs of the experimental and control subjects were imported and scored blindly by this investigator using the TAT protocols that had been coded by an independent investigator. The stories of these records proved to be more difficult to analyze and score than TAT stories written by the subject himself according to Arnold's special instructions that insist on stories with plot and outcome. At the Counseling and Psychotherapy Research Center of the University of Chicago the TAT was administered individually, the subject giving his stories by talking through a microphone in the presence of the test administrator who often encouraged him to use the five minutes allotted to him. This procedure produces long, repetitious, rambling records that fill from seven to fifteen typewritten pages after they were transcribed. It took the rater four hours on the average to import and score each one of the records. Six to twelve months later forty TAT records selected at random were rescored by the rater. His intrajudge reliability was .75. The correlation was modest but fairly adequate. However, it was not possible to assign the clients names to these TAT records with any certainty because the coding was often in error. Hence it was decided to rescore all 156 TATs, using this time the protocols found in the individual folders of the subjects in order to know for sure which were experimental and which control, and which belonged to the pre-therapy, post-therapy and follow-up periods of testing. Obviously, this second scoring could not be blind except that the rating of the experimental subjects by their therapists was not known to the scorers. Under these circumstances, the interjudge reliability of the rater had to be tested by a blind analysis of random records by two

independent raters trained in the SSA method. In order to save some time and labor, the reliability of the scores of the imports of the first eleven stories of each record versus the scores derived from all twenty stories of 100 TAT records was explored. The score from twenty imports was found to be a bit lower than the scores from the first eleven imports of the records, perhaps because of the ambiguity of the cards of the second set of Murray's TAT, but the rank order correlation between the two scores was .956. An independent rater scored fifteen TAT records taken at random, and his scores from the first eleven and all twenty stories were compared. His scores from twenty stories was also slightly lower than his scores derived from the first eleven stories. However, the rank order correlation of his scores was .97. Since these two independent correlations of the scores of 11 and 20 stories of these records were unusually high, it was decided to use only the first eleven stories of the available TAT records in the interjudge reliability study and in the task of rescoreing all the records.

For the interjudge reliability study 18 TAT records of the experimental and 18 of the control subjects -- six from pre-therapy, six from post-therapy and six from follow-up-- were chosen by random numbers. All identifying notes were erased from the records leaving only the age, sex and status of the subject as student or nonstudent, and Xerox copies were made for each one of the independent raters. Though the three raters had been trained and had experience in the method of story sequence analysis, they had some sessions of proximate training and supervision with Dr. Arnold. Ten TAT records from another but similar block in the library of

The Psychotherapy and Research Center of the University of Chicago were analyzed and scored by the three raters and their scores were compared with the analysis done by Dr. Arnold. Differences of imports and scores were discussed with her. After this common practice and tuning up, each rater analyzed and scored separately and independently the 36 TAT records of the reliability study. The two independent raters also found these verbal TAT protocols more difficult to import and score than the written TAT records they had scored in other studies, and spent two hours on the average in the analysis and scoring of each record. The scores of the imports of each record were then added algebraically and transformed into a Motivation Index (MI) following Arnold's ratios (Arnold 1962, p. 146). It was also observed that each rater had his own focus of evaluation, one scoring consistently higher or lower than the other two raters. Some of the scoring categories varied from rater to rater and so did the sequence; their MIs seldom agreed exactly. However, the rank order correlations of the three raters were as follows: between this investigator and rater M, the most experienced in this method, it was .84; between the investigator and rater D it was .70, and between raters M and D it was .66. The interclass reliability of the three raters calculated by the analysis of variance suggested by Guilford was .92 (Guilford, 1954, p. 395). Since the intrajudge and interjudge reliability of this investigator was adequately high, it was not necessary, nor was it feasible, to ask the other raters to import and score all the TAT records of this study. Hence, to preserve consistency of focus of evaluation and uniformity of scoring, ~~of~~ only the scores of this investigator obtained from his rescoring the

first eleven stories of each record, transformed into a Motivation Index, will be used to test the hypotheses of the study in the comparison of the experimental and control subjects as well as of the success and failure cases at the beginning, end, and follow-up periods of therapy and testing. These same scores will be used in the comparison of the MIs of the subjects with the scores they obtained on the other psychological tests they took together with the TAT. To combine the scores of the other two raters, when available, with the scores of the investigator would only serve to confuse the different foci of evaluation taken by each rather, and to average differences in the scores of some subjects which might be important to discriminate adequately the different periods of therapy and of testing.

CHAPTER IV

RESULTS

The rating scale of the Story Sequence Analysis has no equal intervals; it can only be considered an ordinal scale, or at most an "ordered metric scale" as Siegel says (Siegel, 1956, p. 76). As mentioned before (see p. 28), the first studies used only two scores for each import, plus or minus. When an intensity score was introduced by using a scale of 1, 2, 3 and 4, with 1 indicating the most negative and 4 the most positive import this left no room for imports which had no outcome (and often no plot) and so could not be scored as either negative or positive. This difficulty was solved by excluding records containing more than two such unscorable stories. However, by eliminating the TAT records of subjects that "solved" the TAT problem just as they solved their other problems in life, namely, by drifting along without facing the issue and taking a stand, the investigators eliminated a particular type of person altogether. This difficulty was solved by Arnold by placing the zero as the middle point in the scale, and by indicating the intensity score more clearly as -1, -2 or +1 and +2 (Arnold, 1962). Unfortunately, zero is seldom used in the SSA where the sequence is the backbone of the method and where a story can be scored in the context of the preceding or following stories. In more than 5500 stories rated by this investigator there were no zero scores; no zero scores were given by any of the independent raters either. When the rater changes from positive to negative or vice versa, he takes an unequal larger step than when he scores 1 or 2 on the same side of the scale. The intrinsic

sic inequality of steps in the rating scale will delimit the statistics used in this study. Only nonparametric methods that depend on ranking order will be used. However, tests like Spearman Rank Order Correlation, Wilcoxon Matched-Pairs Signed Ranks and Mann-Whitney U Test have been shown to possess high discriminative power and are independent of the requirements of interval scales and the assumptions of parametric statistics. The critical ratio for the rejection of the null hypothesis will be set at the .05 level of significance as in most researches in the behavioral sciences. Results indicating a strong tendency in a particular direction will be indicated by giving probabilities below the .10 level. When there are ties in the ranking order, the formulas suggested by Siegel to correct for this effect will be used (Siegel, 1956).

1. MEASUREMENT OF MOTIVATIONAL CHANGE DURING THERAPY

The main purpose of the study was to measure the change in motivational attitudes effected by client-centered therapy, and to test the stability of this change after therapy. The change was measured by comparing the MI scores of the experimental Ss from pre-therapy to post-therapy and from pre-therapy to follow-up, six to twelve months after the termination of therapy. The stability of the change was measured by comparing the MI scores from post-therapy to follow-up.

To eliminate alternative hypotheses, the change in motivational attitudes presumably effected by therapy in the experimental group will be compared with the change of the control subjects matched with them in sex, education, age, socio-economic status and status as students and nonstudents. The control Ss took the TAT at approximately the same time interval as the

experimental subjects but without taking therapy. The time interval from pre-therapy to post-therapy varied with each experimental subject according to the length of therapy. The range was 22 days to 1176 days, the median being 170 days. The median of the time interval from post-therapy to follow-up was 203 days. The median time interval between the first and second TAT for the control subjects was 178 days, and from the second to the third TAT was 219 days. Though the control subjects did not take therapy, their MI scores will also be called for clarity of the comparison, pre-, post-, and follow-up MI scores the same as for the experimental Ss. It was assumed that the Motivation Index of the control Ss would remain constant or vary randomly throughout the three testing periods. It was also assumed that whatever other environmental influence affected their scores would affect equally the experimental subjects matched with them, and cancel out.

The basic data for our study are in Table 1 for the experimental Ss and in Table 2 for the controls. Four control subjects were dropped

Table 1

Table 2

because it was known that they had refused to come for their second and third tests, and no comparisons were possible. There are other gaps both in the experimental and control groups. The subjects left in the Table were known to have taken the TAT, but their TAT records were scrambled, as explained before; four TAT records were not transcribed at all. The investigator was able to transcribe two tapes and to identify some TAT records from the coded list but some gaps still remained. For this reason, the number of Ss in various calculations will vary with the material that is available.

TABLE 1

Basic data of experimental subjects

Code name	Code number	Sex	Age	Stud. Nonst.	PRE MI	POST MI	F-UP MI	Rating by C.
Bafo	1	F	40	NS	68	114	132	8
Bevi	2	F	29	NS	77	86	105	8
Biva	3	F	21	S	32	86	73	6
Biwi	4	M	27	S	77	64	77	4
Dezo	5	M	24	NS	82	73	64	5
Dura	6	M	31	NS	36	23	none	1
Faja	7	M	28	NS	91	100	95	6
Fimp	8	M	24	S	18	27	5	6
Jidu	9	F	37	NS	73	86	91	6
Jupa	10	F	21	S	68	95	100	7
Keju	11	F	33	NS	123	164	159	7
Nafe	12	M	26	S	95	109	105	6
Nemi	13	M	21	S	45	32	none	5
Pebo	14	M	30	S	50	55	55	6
Pudi	15	M	22	NS	59	41	36	4
Rafo	16	M	28	NS	59	82	95	6
Reba	17	M	26	S	55	50	18	7
Tazi	18	F	36	S	127	136	123	5
Vapa	19	F	23	S	68	82	77	7
Vesi	20	F	26	NS	77	155	159	7
Yima	21	F	28	NS	50	159	55	4
Zoti	22	M	27	NS	45	41	68	5
Dajo	23	M	28	S	73	77	68	6
Fato	24	M	24	S	73	59	50	2
Kibo	25	M	30	S	50	77	50	7
Pija	26	F	22	S	55	68	68	9
Zive	27	M	21	S	105	91	95	7
Voli	28	M	24	NS	64	55	36	2
Lozo	29	M	28	NS	73	55	36	3
Leme	30	M	26	S	50	none	59	8
Md		M 19 F 11	26.8	NS 14 S 16	67.8	77.0	68.4	

Median of success Ss (≥ 6).....67.8 85.8 84.0Median of failure Ss (≤ 5).....61.5 55.0 52.5

TABLE 2

Basic data of control subjects

Code name	Code number	Sex	Age	Stud-Nonst.	PRE MI	POST MI	F-UP MI
Beant	101	F	44	NS	155	155	123
Asmpre	102	F	29	NS	77	73	68
Alele	103	F	19	S	59	55	Tape inaud.
Beste	104	M	30	NS	141	159	Tape inaud.
Aroby	105	M	24	S	45	32	36
Acopre	106	M	29	S	59	64	73
Agest	107	M	25	S	86	82	none
Afesa	108	M	23	S	59	none	none
Avepa	109	F	31	NS	132	127	none
Anepre	110	F	24	NS	95	82	64
Ajopre	111	F	27	NS	155	123	none
Awepre	112	M	30	S	77	95	82
Asmse	113	M	21	S	68	82	86
Becur	114	F	30	NS	132	127	105
Acobre	115	M	22	S	27	23	36
Astpre	116	M	30	S	68	45	50
Bewol	117	M	23	S	none	41	64
Begen	118	F	35	NS	91	82	91
Amcle	119	F	25	S	14	23	32
Apipa	120	F	27	NS	95	86	none
Afypre	121	F	26	NS	55	100	114
Ashube	122	M	28	S	64	59	73
Betur	123	F	26	NS	23	27	27
Bepen	124	F	33	NS	127	91	91
Betho	125	F	45	NS	86	68	91
Betar	126	M	26	S	95	64	36
Medians		F 14 M 12	27.0	S 13 NS 13	77.0	77.2	70.2

The main hypotheses are given in the Introduction of this study. Here they will be repeated one by one as the relevant statistical evidence is presented.

Hypothesis 1: Experimental Ss who received therapy for more than six sessions will show significant improvement in their motivational attitudes at the end and follow-up of therapy, i.e. will have significantly higher MI at post-therapy and follow-up than at pre-therapy. This is presumably attributed to therapy.

Three measurements were made: from pre-therapy to post-therapy, from pre-therapy to follow-up and from post-therapy to follow-up because Rogers claims that therapy starts a process of change that is better shown some time after therapy is completed. (Rogers, C. & Dymond, R., 1954). Others argue that the effects reported in some therapy studies are of the "hello-goodbye" type, and do not last after therapy (Hathaway, 1948). The stability of the changes effected during therapy will be tested by comparing the post-therapy and follow-up MI scores.

The first hypothesis was broken down into three parts which were tested separately:

1. Experimental Ss show a significant increase in their MI scores from pre-therapy to post-therapy. This was tested by the Wilcoxon Matched-Pairs Signed-Ranks Test. This statistic was used because it takes into account the amount as well as the direction of change and is therefore more powerful than the Sign Test. The Wilcoxon Test gives a T of 135 which with 25 pairs is significant at the .05 level in a one-tailed test ($z=1.79$ see Table 3). Thus Hypothesis 1 is verified.

TABLE 3

Wilcoxon Matched-Pairs Signed-Ranks Tests

Subjects	Pairs compared	Points of comparison	z from obtained T	Probability (one-tail)	Direction
Experimental Ss	25	Pre to Post	1.79	p = .04	Increase
Experimental Ss	25	Pre to Fup	1.16	p > .10	Increase
Experimental Ss	25	Post to Fup	1.17	p > .10	Decrease
Control Ss	19	Pre to Post	2.01	p = .02	Decrease
Control Ss	17	Pre to Fup	.72	NS	
Control Ss	17	Post to Fup	.30	NS	
Success Exp. Ss	16	Pre to Post	3.17	p = .002	Increase
Success Exp. Ss	15	Pre to Fup	2.11	p = .02	Increase
Success Exp. Ss	15	Post to Fup	.28	NS	
Failure Exp. Ss	11	Pre to Post	1.68	p = .05	Decrease
Failure Exp. Ss	11	Pre to Fup	1.61	p = .05	Decrease
Failure Exp. Ss	11	Post to Fup	1.48	p = .07	Decrease
Female Exp. Ss	11	Pre to Post	2.93	p = .002	Increase
Female Exp. Ss	11	Pre to Fup	2.84	p = .002	Increase
Female Exp. Ss	11	Post to Fup	.25	NS	
Male Exp. Ss	18	Pre to Post	.76	NS	
Male Exp. Ss	17	Pre to Fup	1.36	p = .09	Decrease
Male Exp. Ss	16	Post to Fup	1.16	p > .10	Decrease
Female control Ss	13	Pre to Post	1.86	p = .03	Decrease
Female control Ss	9	Pre to Fup	1.06	p > .10	Decrease
Female control Ss	8	Post to Fup	.77	NS	
Male control Ss	10	Pre to Post	.30	NS	
Male control Ss	8	Pre to Fup	.91	NS	
Male control Ss	9	Post to Fup	1.75	p = .04	Increase
Female success Ss	9	Pre to Post	2.63	p = .004	Increase
Female success Ss	9	Pre to Fup	2.52	p = .01	Increase
Female success Ss	9	Post to Fup	1.39	p = .08	Increase
Male success Ss	9	Pre to Post	1.68	p = .05	Increase
Male success Ss	9	Pre to Fup	.01	NS	
Male success Ss	7	Post to Fup	1.29	p > .10	Decrease

2. The experimental subjects show a motivational increase from pre-therapy to follow-up, i.e., their MI scores at follow-up are higher than at pre-therapy. The Wilcoxon Test gives a T of 129.5 for twenty-five pairs of scores. It is not significant at the .05 level ($z=1.16$). Rogers' claim of continuous progress after therapy was not confirmed as far as MI scores were concerned.

3. The experimental subjects maintain the therapy change from post-therapy to follow-up. The average time from post-therapy to follow-up was 203 days, and the average time of therapy was 170 days. Though the Wilcoxon Test shows a slight decrease in the MI scores of the experimental subjects from post-therapy to follow-up, it is not significant at the .05 level ($z=1.17$); hence it can be concluded that the positive change achieved during therapy is maintained for a period of time approximately equal to the time of therapy. (See Table 3).

Hypothesis II: The control Ss who took the TAT at approximately equal time intervals as the experimental Ss but did not take therapy will not show any significant improvement in their motivational attitudes from the first to the second or from the first to the third tests, i.e., their MI scores at post-therapy and follow-up will not be significantly higher than at pre-therapy.

It was assumed that the MI scores of the control Ss would remain constant or vary at random since they had no therapy. This hypothesis was also tested in three parts by the Wilcoxon Signed-Ranks Test:

1. Out of nineteen control Ss who showed change in their MI scores from pre-therapy to post-therapy, fifteen decreased their MI scores at post-

therapy. The Wilcoxon Test gives a T of 45 ($z=2.01$, $p=.02$). Instead of increasing their MI scores from pre-therapy to post-therapy, the control Ss showed a significant decrease.

2. From pre-therapy to follow-up there were seventeen matched pairs of scores with positive and negative changes. The changes were random and not significant.

3. From post-therapy to follow-up eleven out of seventeen control Ss made a slight increase in their MI scores but the Wilcoxon T is 70 and is not significant at the .05 level.

These results confirm the second hypothesis, namely, control Ss without therapy do not show a significant increase in motivation during a time interval approximately equal to the time of therapy and follow-up of the experimental Ss. On the contrary, they showed a significant decrease in their MI scores from pre-therapy to post-therapy. As a possible explanation for this decrease it can be speculated that the TAT has a definite impact on the testee. Taking the TAT is beneficial if it is followed by therapy or at least by a brief discussion of the problems revealed in the TAT with the test administrator as shown in Bettelheim's and Morton's experiments reported in the review of related literature (Bettelheim, 1947; Morton, 1955). It could be prejudicial if the testee does not have the advantage of therapy or a brief discussion of his problems with the test interpreter after taking the TAT, as was the case with the control Ss in our study. However, this is only speculation that needs study. The results of the first and second hypotheses can be seen diagrammatically represented in Figure 1 using the medians of the MI scores at pre-therapy, post-therapy and

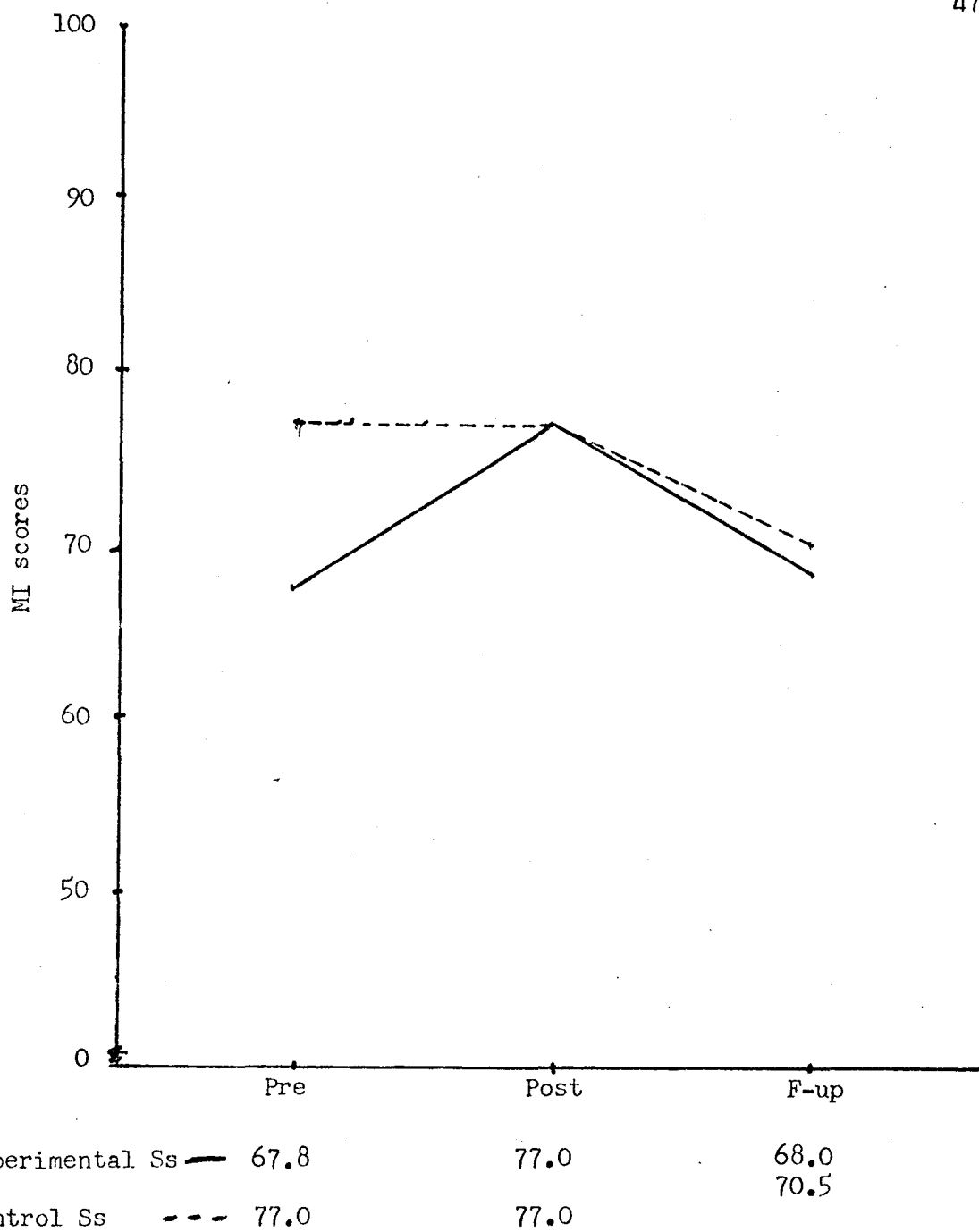


Figure 1. Median MI scores of experimental and control subjects.

follow-up as points of comparison rather means as more appropriate with an ordinal scale. The figures may differ slightly from the results reported in the test due to the gaps in the data and the nature of the Wilcoxon Test. Like the Sign Test, the Wilcoxon Test eliminates data where the pairs of scores are incomplete or show no change. The figures were drawn from all the data available at each point of comparison and are independent of the limitations of the Wilcoxon Test.

Research on the outcome of psychotherapy often shows inconclusive results because therapy groups are ordinarily composed of two quite different subgroups, namely, the success and failure cases. Therapy has a strong impact on the subject. This is shown as improvement in success cases or as deterioration where therapy proves to be a failure by fault either of the patient or the therapist or any other cause. The counterbalancing influences of the success and failure cases may account for the ambiguity of results often shown by studies on the outcome of therapy. In order to clear up any possible ambiguity in the results of this study the experimental group was divided into success and failure subgroups on the basis of the therapists' ratings at the end of therapy. Subjects who received a rating of 6 or more on a 9-point scale will form the success subgroup. The failure subgroup will be formed by patients who received a score of 5 or less at the end of therapy. On this basis, there are nineteen success and eleven failure cases in our sample. Both will be compared with each other and with the control group at pre-therapy, post-therapy and follow-up.

Hypothesis III: Experimental Ss rated as success cases by their therapist at the end of therapy will show a significant increase in moti-

vation, i.e., they will have significantly higher MI scores at post-therapy and follow-up than at pre-therapy.

This hypothesis was also tested by the Wilcoxon Test in three separate parts:

1. From pre-therapy to post-therapy there were eighteen pairs of MI scores that show change: sixteen show a positive change or increase, and only two show a negative change or decrease. The Wilcoxon Test gives a T of 12.5 with a z of 3.17. This has a probability of .002 indicating that Ss rated as success cases had a very significant increase in their MI scores from pre-therapy to post-therapy (See Table 3).

2. From pre-therapy to follow-up fifteen Ss increased their MI scores four had some decrease. The Wilcoxon T is 37, significant at the .05 level ($z=2.11$, $p=.02$).

3. From post-therapy to follow-up the change is random. The Wilcoxon test is not significant, thus showing that the MI at follow-up is not significantly different from the MI at post-therapy. Though the success Ss did not increase their MI scores after therapy, they maintained the motivational development acquired during therapy for at least an equal period of time. The third hypothesis is confirmed.

Hypothesis IV: Patients rated as failure cases by their therapists will show no significant increase in motivation, i.e., they will have no significantly higher MI scores at post-therapy and follow-up than at pre-therapy.

This hypothesis will be tested also in three parts:

1. From pre-therapy to post-therapy nine subjects decreased their MI scores and only two showed an increase. The resulting T is 14.0 and is

significant at the .05 level ($z=1.68$, See Table 3). Instead of increasing their MI scores, they showed a decrease as was observed with the control Ss.

2. From pre-therapy to follow-up there is also a decrease in their MI scores. The Wilcoxon Test gives a T of 6.5 which is significant at the .05 level (See Table 3).

3. From post-therapy to follow-up, the Wilcoxon Test shows a decrease in the MI scores but the resulting T has only a probability of .07 (See Table 3).

Hypothesis IV is significantly confirmed. What is not yet clear is why the failure subjects have a definite decrease rather than maintaining their MI scores. Again we can only speculate that therapy is not indifferent: it either helps the subjects to improve their motivational attitudes when it is successful or it brings out the negative attitudes more openly when it is a failure. The graphic description of these results can be seen in Figure 2 with median scores at pre-therapy, post-therapy and follow-up as points of comparison.

Figure 2

Arnold claims that the SSA reveals the habitual motivation of the subject. In the comparisons made so far, the SSA has proved its ability to detect changes--positive or negative--that follow therapy. Subjects who were rated as success in therapy (like high achievers in school and business) show that when they set out to work on something they make a success of it by personal effort and the use of adequate means, while subjects rated as

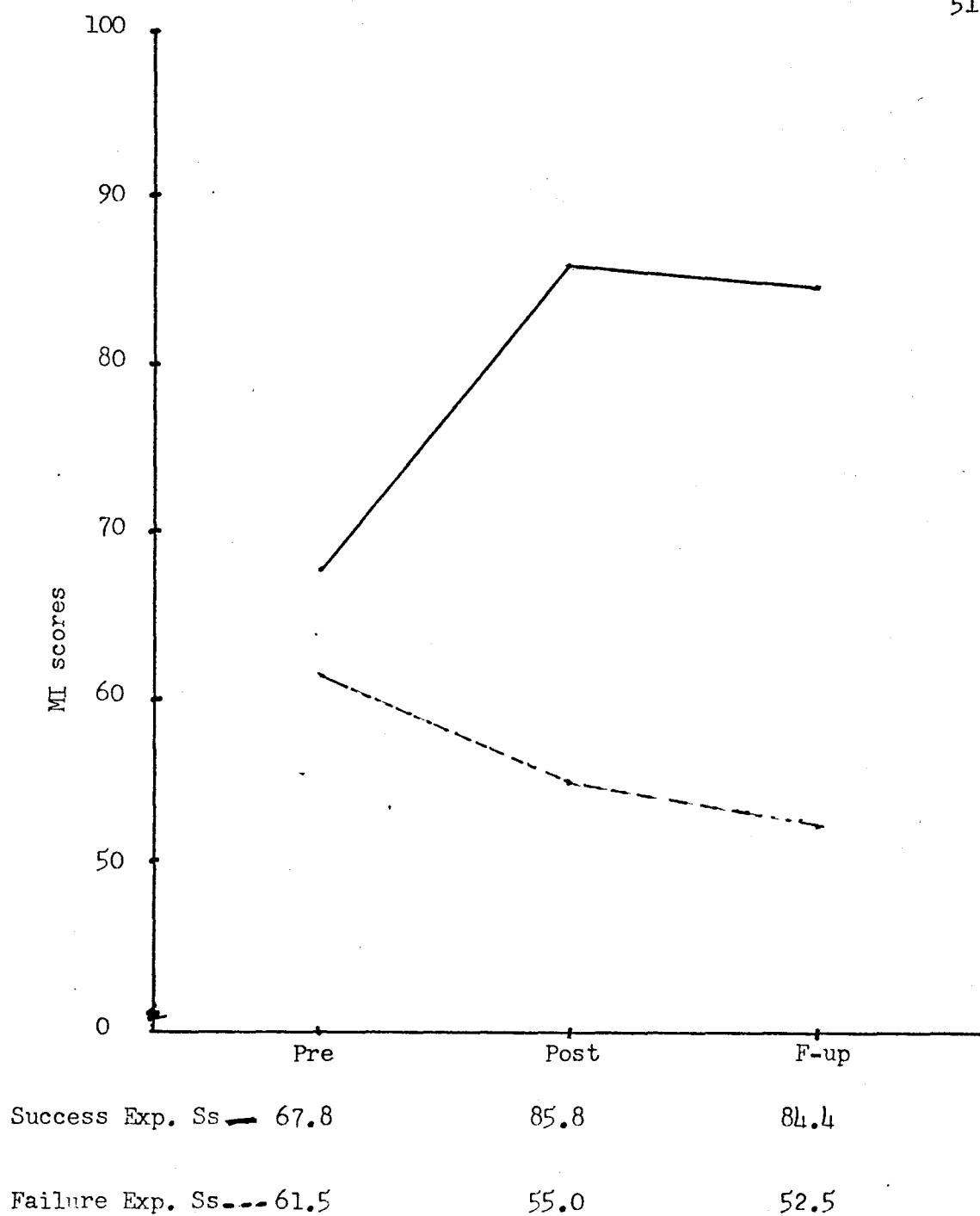


Figure 2. Median MI scores of success and failure experimental subjects.

failure cases presumably take therapy as they attack other problems in life, namely, passively, hoping that somebody else will do the job for them and perhaps get frustrated when the therapist, especially of the client-centered orientation, refuses to do this for them. If such is the situation it can be hypothesized that success subjects will show higher motivation than failure cases even at pre-therapy.

Hypothesis V: "Success" experimental Ss will be better motivated than failure cases at the beginning of therapy, i.e., they will have higher MI scores at pre-therapy than failure cases.

It has been shown that success therapy subjects significantly increased their MI scores at post-therapy and follow-up. Figure 2 shows that the failure cases had lower MI scores at pre-therapy as well as at post-therapy and follow-up as expected according to Hypothesis V. However, it was tested more strictly by comparing the success group with the failure group, at pre-therapy, post-therapy and follow-up by the Mann-Whitney U Test which is not restricted by the requirements of parametric tests and has a power-efficiency of 95.5 percent as compared with the t test (Siegel, 1956, p. 126). The resulting U's were 97, 48.5 and 48.5 at the three points of comparison with probabilities of .37, .01 and .02. This indicates that the experimental Ss rated as "success" therapy cases had a significantly higher MI than the failure cases at post-therapy and follow-up, but not at pre-therapy. Hypothesis V was not confirmed. This seems to indicate that the subjects rated as success at the end of therapy significantly increased their MI scores at post-therapy and follow-up as a result of therapy and not of motivational differences they might have had before therapy.

TABLE 4
Mann-Whitney U Tests

Subjects	Point of comparison	z from obtained U	Probability (two-tailed test)
Exp. and control Ss	Pre-therapy	1.75	p = .08
Exp. and control Ss	Post-therapy	.21	NS
Exp. and control Ss	Follow-up	.13	NS
Exp. success & control Ss	Pre-therapy	1.40	p > .10
Exp. success & control Ss	Post-therapy	3.16	p = .002
Exp. success & control Ss	Follow-up	1.36	p > .10
Exp. success & failure Ss	Pre-therapy	.90	NS
Exp. success & failure Ss	Post-therapy	2.60	p = .01
Exp. success & failure Ss	Follow-up	2.20	p = .02
Exp. failure & control Ss	Pre-therapy	1.44	p > .10
Exp. failure & control Ss	Post-therapy	1.24	p > .10
Exp. failure & control Ss	Follow-up	.77	NS
Male & female Exp. Ss	Pre-therapy	.95	NS
Male & female Exp. Ss	Post-therapy	3.47	p < .001
Male & female Exp. Ss	Follow-up	3.16	p = .002
Exp. & control female Ss	Pre-therapy	1.99	p = .05
Exp. & control female Ss	Post-therapy	1.34	p > .10
Exp. & control female Ss	Follow-up	1.27	p > .10
Male & female control Ss	Pre-therapy	2.56	p = .01
Male & female control Ss	Post-therapy	1.86	p = .06
Male & female control Ss	Follow-up	1.76	p = .08
Female suc. & female cont. Ss	Pre-therapy	2.05	p = .04
Female suc. & female cont. Ss	Post-therapy	.98	NS
Female suc. & female cont. Ss	Follow-up	1.35	p > .10
Male success & female suc. Ss	Pre-therapy	.75	NS
Male success & female suc. Ss	Post-therapy	1.77	p = .08
Male success & female suc. Ss	Follow-up	2.21	p = .03
Male success & male failure Ss	Pre-therapy	.28	NS
Male success & male failure Ss	Post-therapy	2.08	p = .04
Male success & male failure Ss	Follow-up	.98	NS

When the whole experimental group (success and failure groups combined) were compared with the controls at the three testing points by means of the Mann-Whitney U test, no significant differences were found at any point, though the experimental group showed a tendency toward lower M.I.s at pretherapy.

Comparing the success group alone with the controls, the following U's were obtained: 190, 179 and 134. The corresponding z's were: 1.40, 3.16 and 1.36 (see Table 4). The differences at pre-therapy and follow-up are not significant at the .05 level, although there is again a slight tendency toward lower M.I. scores in the success group. Since the differences at post-therapy are significant at the .01 level, the suggestion that the differences are the effect of therapy is confirmed. When compared with the controls, the failure experimental Ss showed no significant differences at pre-therapy, post-therapy or follow-up, though there was again a tendency toward lower M.I.s (see Table 4, see Figures 3 and 4).

Though there are no significant differences between experimental and control subjects at pretherapy, there seems to be a slight tendency for experimental Ss in general and for success subjects in particular to have lower MIs. A similar phenomenon has been repeatedly observed by other investigators doing research on the outcome and process of therapy (Butler & Haigh, 1954; Dymonds, 1954; Rogers, 1954). Butler and Haigh speculated that this may be due to the low self-esteem and the low self-concept of subjects applying for therapy. However, in Rogers' study (1954) as well as in the present study this explanation does not seem adequate since the rating on the Emotional Maturity Scale was not done by the subjects and the MI does not depend on self-esteem or self-concept.

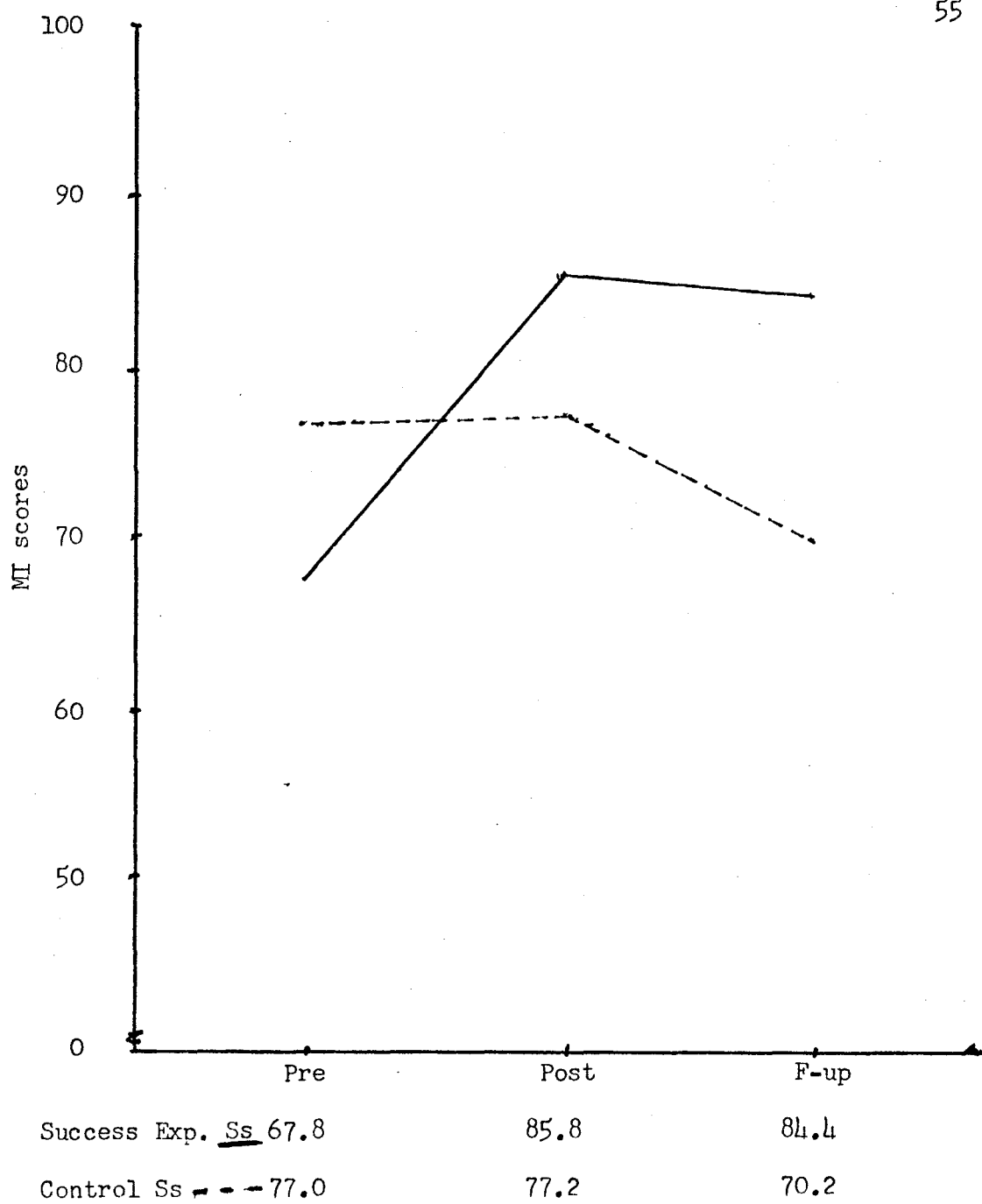


Figure 3. Median MI scores of success experimental and control subjects

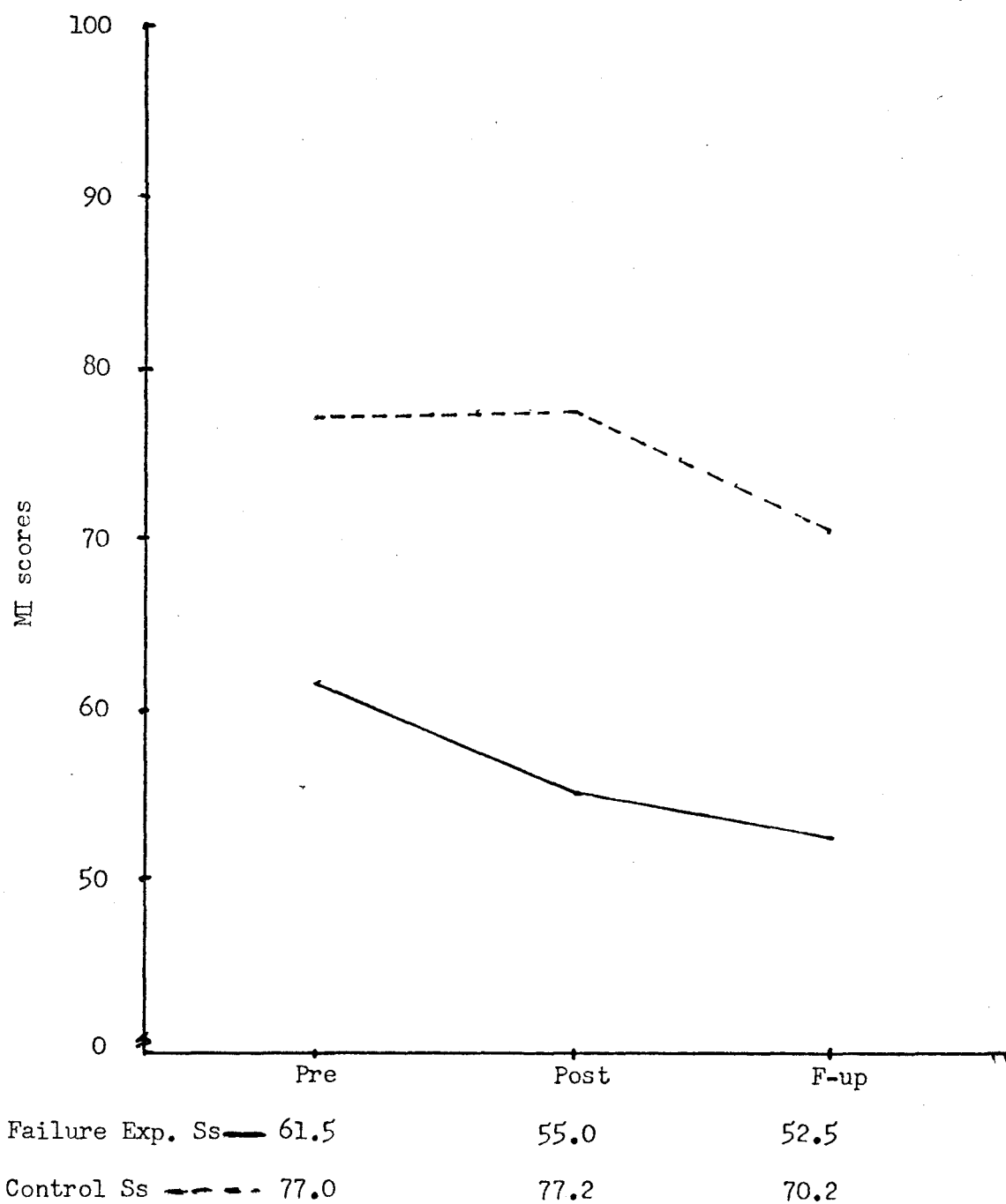


Figure 4. Median MI scores of failure experimental and control subjects

II. COMPARISON OF THE STORY SEQUENCE ANALYSIS WITH OTHER PSYCHOLOGICAL TESTS

Story Sequence Analysis has been applied here for the first time to evaluate the outcome of client-centered therapy. Though applied to therapy here for the first time, the SSA has been shown to be adequate to discriminate reliably and consistently between experimental and control subjects as well as between success and failure subjects so rated by client-centered therapists who value the internal focus of evaluation of the client more than any external criterion. It was of great interest, therefore, to examine further the validity of the SSA as a measure of the outcome of therapy by comparing the MI scores with the scores of other tests used for the evaluation of client-centered therapy. Five tests were selected for this comparison: the Self-Ideal Self Q-Sort, Willoughby Emotional Maturity Scale, the Counselor's Rating, the Personal Integration Rating and the TAT clinically analyzed and scored by an independent diagnostician following a different method of TAT analysis.

A. The Motivation Index and Self-Ideal Self Q-Sort

Butler and Haigh developed the Self-Ideal Self Q-Sort from Stephen's Q-Technique, Rogers' idea of the Self-Concept and their own definition of the Ideal Self. Rogers defined the Self-Concept as an "organized, fluid but consistent, conceptual pattern of the characteristics of the 'I' and 'me' which are admissible to awareness, together with the values attached to these concepts". (Rogers, 1951, p. 498). The authors defined the Ideal Self as "the organized conceptual patterns of characteristics and emotional states which the individual consciously holds as desirable (or undesirable)

for himself". (Butler & Haigh, 1954, p. 56). One hundred self-referent statements taken from actual therapy interviews are rated by the S as least or most like himself on an 11-point scale approximating a normal distribution. This constitutes the Q-Sort for the Self-Concept. For the Ideal Self the S rates the same statements on a forced 11-point distribution as least or most desirable for himself. This constitutes the Ideal-Self. These two Q-Sorts are then correlated. Butler and Haigh hypothesized that the correlation between the two Q-Sorts would be low at pre-therapy due to the low self-esteem and low self-concept of the client at pre-therapy, but that this correlation would increase after therapy and at follow-up by the modification and upgrading of the self-concept during therapy. In their study they showed that the Self-Ideal Self correlations of subjects seeking therapy and especially of those rated as success increased significantly at post-therapy and follow-up (Butler & Haigh, 1954).

The Story Sequence Analysis purports to reveal the organized patterns of motivational attitudes of the subject by means of the imports of the TAT stories. The subject may or may not be aware of them, but he reveals them in the plot and outcome of his stories. Since the Q-Sort depends on the conscious self-rating of the subject as he is and as he would like to be, it may be distorted by the defensive sortings of the subject. Since the Q-Sort depends on the subjective focus of evaluation while the SSA depends on the less conscious and more objective self-revelation through the imports of the TAT stories, it would appear that they would show little correlation. The correlations of the MI scores and the Q-Sorts of the experimental and control subjects are given in Table 5. They are not significant at the .05

TABLE 5

Rank order correlations of MI scores and Self-Ideal Self r's

	Pre-therapy MI	Post-therapy MI	Follow-up MI
Self-Ideal r's			
Experimental Ss N=25	-.14	.31	.32
Control Ss N=16	.03	No data	-.15

level.

The experimental group was divided into success and failure subgroups, and their MI's were compared with their Self-Ideal Self r's at pre-therapy, post-therapy and follow-up. They are given in Table 6. None of the correlations are significant at the .05 level. It would seem that the self-concept and the self-ideal with ratings of the Ss is not related to their motivational attitudes as expressed in TAT stories.

B. The Motivation Index and the Emotional Maturity Scale

The Willoughby Emotional Maturity Scale combines the internal focus of evaluation of the subject and the perspective of outside observers. It consists of 69 statements descriptive of the emotional maturity of a person as rated previously by a group of clinicians on a 9-point scale. The subject checks the items that fit him; this is the subjective view of himself. Two of his friends who are supposed to know him well check the items that in their opinion are descriptive of the subject; this is the objective view of the subject. Rogers used this test to evaluate the outcome of client-centered therapy. He found that the client and the observers are consistent in their ratings but that their different views have little or no correlation. He also found that the subjects rated as success after therapy tend to devalue themselves. Their E-M self-ratings are lower than the ratings given them by observers, especially at pre-therapy. On the other hand, the self-ratings of failure Ss are higher than the ratings given them by outside observers. He found also that the self-ratings of the success clients as well as the ratings given them by their friends increase from pre-therapy to post-therapy and from pre-therapy to follow-up. The failure Ss, on the

TABLE 6

Rho correlations between MI's and Self-Ideal Self r's
of success and failure experimental Ss

	Pre-therapy MI	Post-therapy MI	Follow-up MI
Self-Ideal r's			
Success Exp. Ss N=17	-.24	.02	.18
Failure Exp. Ss N=8	.12	.61	.48

other hand, rate themselves very high at post-therapy and follow-up, but their friends rate them lower at post-therapy and follow-up than at pre-therapy (Rogers, 1954).

The MI scores of the experimental Ss as a whole and of the success and failure subgroups were correlated with their self-ratings on the E-M scale and with the ratings given them by outside observers on the same scale. (The ratings of the observers were averaged).

See Tables 7, 8 and 9

There is only one correlation significant at the .05 level, namely, the correlation between MI and E-M ratings of the experimental Ss (.33, Table 7), and that is so low that it cannot be assumed that there is any relationship between motivational attitudes and emotional maturity self-ratings. The correlation of .62 ($p=.05$, Table 9) of failure Ss at post-therapy seems to indicate that this group is responsible for the correlation observed among the experimental Ss as a whole. It would seem that the failure Ss have at least learned insight into their emotions. (The control Ss did not take the E-M).

C. The Motivation Index and the Counselors' Ratings

At the end of therapy the Counselors rated the outcome of therapy on a 9-point scale, with score 1 indicating little or no improvement and score 9 maximal improvement. This measure has been found to correlate well with other psychological tests at post-therapy and follow-up (Rogers & Dymond, 1954). The Counselors' rating of the clients at the end of therapy was correlated with the MI scores the clients obtained at post-therapy and

TABLE 7

Rho correlations between the MI scores and the E-M ratings of Exp. Ss

	Pre-therapy MI	Post-therapy MI	Follow-up MI
E-M ratings by clients N=26	.03	.33*	.29
E-M ratings by observers N=23	-.01	.07	.30

* Significant at the .05 level

TABLE 8

Rho correlations between MI scores and E-M ratings of
success experimental Ss

	Pre-therapy MI	Post-therapy MI	Follow-up MI
E-M ratings by clients N=19	-.30	.09	.20
E-M ratings by observers N=16	-.32	.20	.09

TABLE 9

Rho correlations between the MI scores and E-M ratings
of Failure Experimental Ss.

	Pre-therapy MI	Post-therapy MI	Follow-up MI
E-M ratings by clients N=9	.38	.62*	.44
E-M ratings by observers N=6	.46	.03	.10

* Prob. = .05

follow-up. The correlations were: .45 at post-therapy and .37 at follow-up. Both are significant at the .05 level. The correlation is not very high; whether it is the MI that is closer to the actual status of the client or the rating of the therapists is impossible to say.

D. The Motivation Index and the Personal Integration Rating

At the end of therapy the Counselors rated (on a 9-point scale) the degree of personal integration that they thought their clients had had at the beginning of therapy and at the time of the rating, i.e., at the end of therapy. The rating depended exclusively on the Counselors' knowledge of their clients obtained through therapy since they had no access to the results of the other tests. Their rating of their clients as they had been at the beginning of therapy could have been easily distorted by defects of memory and personal bias, and for this reason no significant correlation was expected between this rating and the MI scores at pre-therapy. It was expected, however, that the Counselors' rating of personal integration at the end of therapy would be more accurate and that it would correlate significantly with the MI scores at post therapy. The obtained correlations were: .19 at pre-therapy and .40 at post-therapy. Only the correlation at post-therapy is significant at the .05 level.

E. The Motivation Index and the TAT rating by an independent clinician

An independent clinician scored the TAT records on a 7-point scale of personal adjustment, 1 indicating minimal and 9 maximal adjustment. Though the rating done by this clinician was intuitive and global, it could be considered more objective than the self-ratings of the clients themselves,

and perhaps less subject to the possible personal involvement of the Counselors. This investigator would expect some correlation between the MI scores and the TAT ratings done by an independent diagnostician. Only the ratings at pre-therapy and follow-up were available for the present comparison. The correlations obtained were: .46 at pre-therapy and .49 at follow-up, both significant at the .05 level.

After the various comparisons made between the MI scores of the experimental and control subjects, and of the success and failure cases with each other and the control subjects, it can be concluded that the four main hypothesis of the study were confirmed. Though empirically derived from studies done with normal people who were not asking for therapy, the Story Sequence Analysis appears to be a reliable and valid measure of therapy by its overall agreement with other tests used to evaluate the outcome of client-centered therapy, and by its precise and clear discrimination between experimental and control subjects, as well as between success and failure cases at pre-therapy, post-therapy and follow-up.

CHAPTER V

DISCUSSION OF RESULTS

The experimental and control subjects were matched by sex, age, socioeconomic level, status as student or nonstudent and time interval between the TAT tests. It is known that these actuarial variables may affect the TAT scores (Murstein, 1962), but it was assumed that whatever influence they might have on the MI scores of the experimental subjects would affect equally the MI scores of the controls, and cancel out. This is considered a sound practice in research. However, it may obscure the real effects of therapy and thus contribute to the inconclusive results often reported in studies on the outcome and process of therapy. In the preceding chapter it was shown that the experimental subjects significantly increased their MI scores from pre-therapy to post-therapy. This result was attributed to therapy because it was not shown by the control subjects who had no therapy. On a closer look, however, it appeared that this increase in MI scores was achieved only by the experimental subjects who had been rated as success cases by their therapists. The failure subjects did not increase their MI scores; on the contrary, they as well as the controls had a significant decrease in MI from pre-therapy to post-therapy. This difference between success and failure cases would have been obscured or reduced if the experimental group had been compared as a whole with the controls. It would seem that most of the results in investigations on the outcome of therapy tend to average the positive and negative changes of the success and failure cases, thus contributing to the inconclusive results reported by different studies. If the positive change of the success subjects outweighs the neg-

ative influence of the failure cases the results are positive changes, as in the present study. In other studies the negative influence of the failure cases might outweigh the positive changes of the success cases, and the results would be negative. This double influence can be seen by comparing Figures 1 and 2.

Since in our sample the female experimental Ss had higher MI scores than the male experimental Ss (See Table 1), it could be asked whether the motivational change reported for the whole experimental group was due to sex rather than to therapy. To clear this up, the influence of sex on the MI scores was investigated in detail.

The Motivation Index and the Sex Variable

Studies done on the sex variable and the outcome of client-centered therapy have reported conflicting results. Cartwright (1955) did not find any significant difference in the outcome of therapy between male and female clients. On the other hand, Seeman (1954) found that significantly more female than male clients had been rated as success therapy cases by their therapists. Dymond (1954) also found that in her sample women increased their TAT ratings at post-therapy and follow-up more than men.

In order to investigate whether sex differences affected the MI scores of the experimental Ss, both subgroups of males and females were compared with each other at pre-therapy, post-therapy and follow-up by the Mann-Whitney U Test. The z's from the obtained U's were: .95 (N.S.) at pre-therapy, 3.47 ($p < .001$) at post-therapy and 3.16 ($p = .001$) at follow-up. (See Table 4.) The results indicate that the male and female Ss did not differ statistically at pre-therapy, but that at post-therapy and follow-up

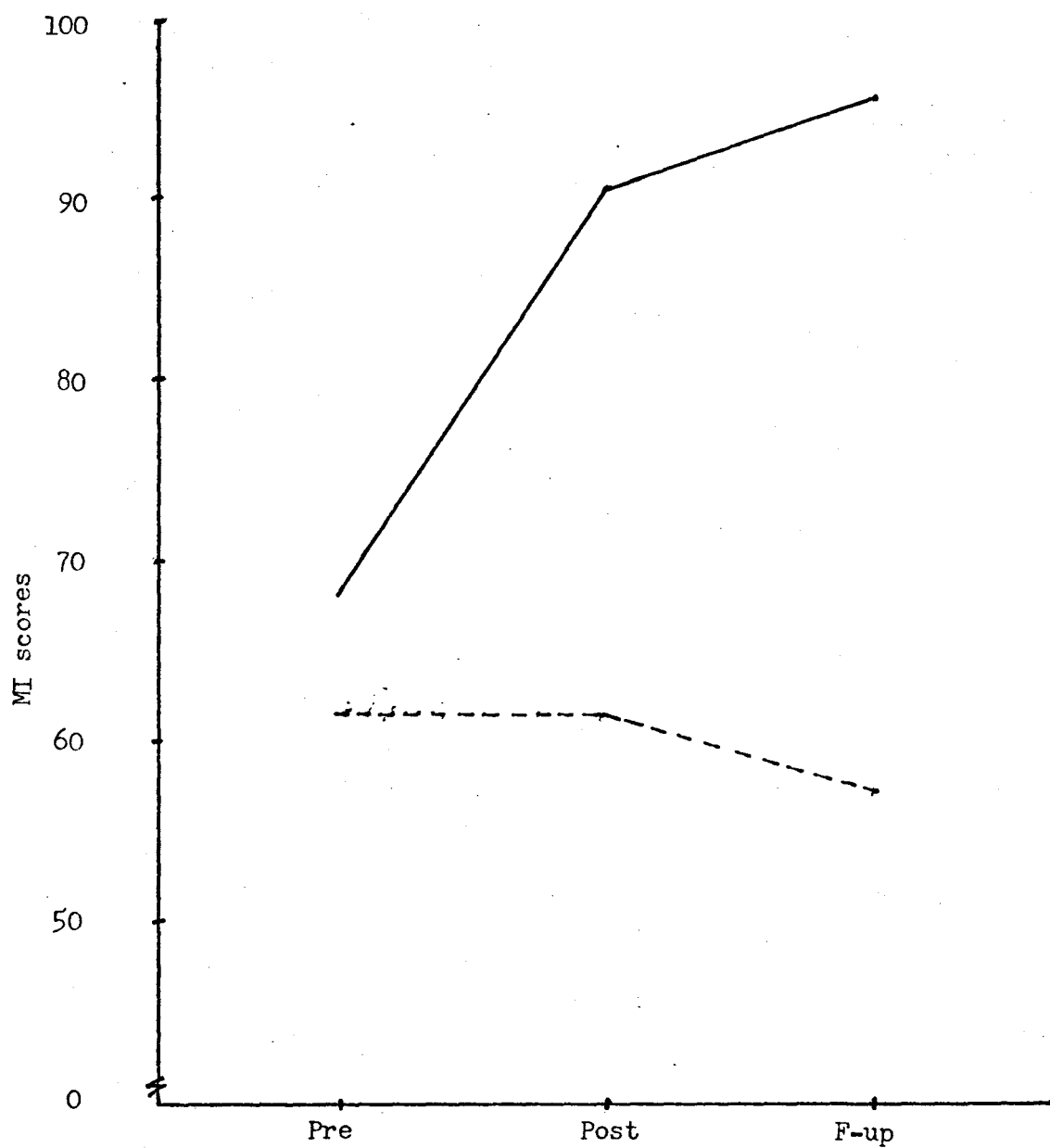
the female Ss differed significantly from the male Ss.

This difference at post-therapy and follow-up was further investigated by examining the change in the MI scores from test to test by the Wilcoxon Signed-Ranks Test. The comparison of the MI scores of the female Ss from pre-therapy to post-therapy, from pre-therapy to follow-up and from post-therapy to follow-up gave the following z 's: 2.93 ($p=.002$), 2.84 ($p=.002$) and .25 (N.S.), indicating that the increase of their MI scores from pre-therapy to post-therapy and from pre-therapy to follow-up was very significant; their change from post-therapy to follow-up was random. (See Table 3)

On the other hand, the male Ss in the experimental group showed a constant though not significant decrease in their MI scores: the increase from pre-therapy to follow-up reaches a probability of .09 in a one-tailed test (See Table 3). It can be concluded, therefore, that though the two subgroups were not statistically different at pre-therapy, the female Ss increased their MI scores significantly during therapy while the male Ss showed a constant decrease. These results are graphically described in Figure 5.

Figure 5

The male and female Ss of the control group were compared also at the three testing points by the Mann-Whitney U Test. At pre-therapy the female control Ss were statistically different from the male control Ss; their MI scores being significantly higher ($z = 2.56$, $p = .01$ in a two tailed test). At post-therapy and follow-up their MI scores are still higher but their differences are not significant at the .05 level. (See Table 4).



Female Exp. Ss —	68.2	90.7	95.5
Male Exp. Ss ---	61.7	61.4	57.0

Figure 5. Median MI scores of female and male experimental subjects.

It was inquired further if the subgroups of the control subjects had any significant increase in the MI scores from test to test. The female Ss had a significant decrease from pre-therapy to post-therapy ($z = 1.86$, $p = .03$. See Table 3). From pre-therapy to follow-up they still showed a decrease in their MI scores but it was not significant. From post-therapy to follow-up the change is random. The male subgroup, on the other hand, had random changes from pre-therapy to post-therapy and from pre-therapy to follow-up. From post-therapy to follow-up they showed an increase in their MI scores significant at the .05 level ($z = 1.75$. See Table 3).

In conclusion, the female Ss of the control group had higher MI scores than the male Ss at pre-therapy; but the MI's decreased from pre-therapy to follow-up. The male Ss had lower scores at the three testing periods but increased their MI scores significantly at follow-up. For the sake of comparison it will be recalled that the female Ss of the experimental group had also higher MI scores than the male Ss at pre-therapy and increased them at post-therapy and follow-up while the male Ss decreased their MI scores at post-therapy and follow-up.

Since the female subjects of the experimental and control groups had higher MI scores than the male Ss they were compared with each other by the Mann-Whitney Test to see if the high scores of the female experimental subjects were due to sex rather than to therapy. The female experimental subjects had significantly lower MI scores than the female controls at pre-therapy ($z = 1.99$ $p = .05$ in a two-tailed test. See Table 4) but they increased them significantly at post-therapy and follow-up while

the female controls decreased theirs. The Mann-Whitney Test does not show any significant difference between the two subgroups at post-therapy and follow-up. See Figure 6.

Figure 6

When only the female experimental Ss who were rated as success cases were compared with the female control Ss, they showed also lower MI scores than the female controls at pre-therapy ($z = 2.05$, $p = .04$ in a two-tailed test). At post-therapy and follow-up they did not show any significant difference (See Table 4, Figure 7.).

Figure 7

The male experimental subjects were also compared with the male control Ss at pre-therapy, post-therapy and follow-up. They were not statistically different at any of the three points of comparison. The differences between the male success Ss and the male Ss of the control group were also insignificant.

In order to explore even further the possible influence of sex on the MI scores of the experimental Ss, the subjects rated as success by their therapists were divided into ten male success and nine female success experimental subjects and were compared with each other by the Mann-Whitney Test. The male and female success subgroups are not statistically different at pre-therapy. At post-therapy the female success Ss had higher MI scores but their difference from the male success subgroups has only a probability of .08 in a two-tailed test ($z = 1.77$). At follow-up the female success Ss were significantly higher and statistically different

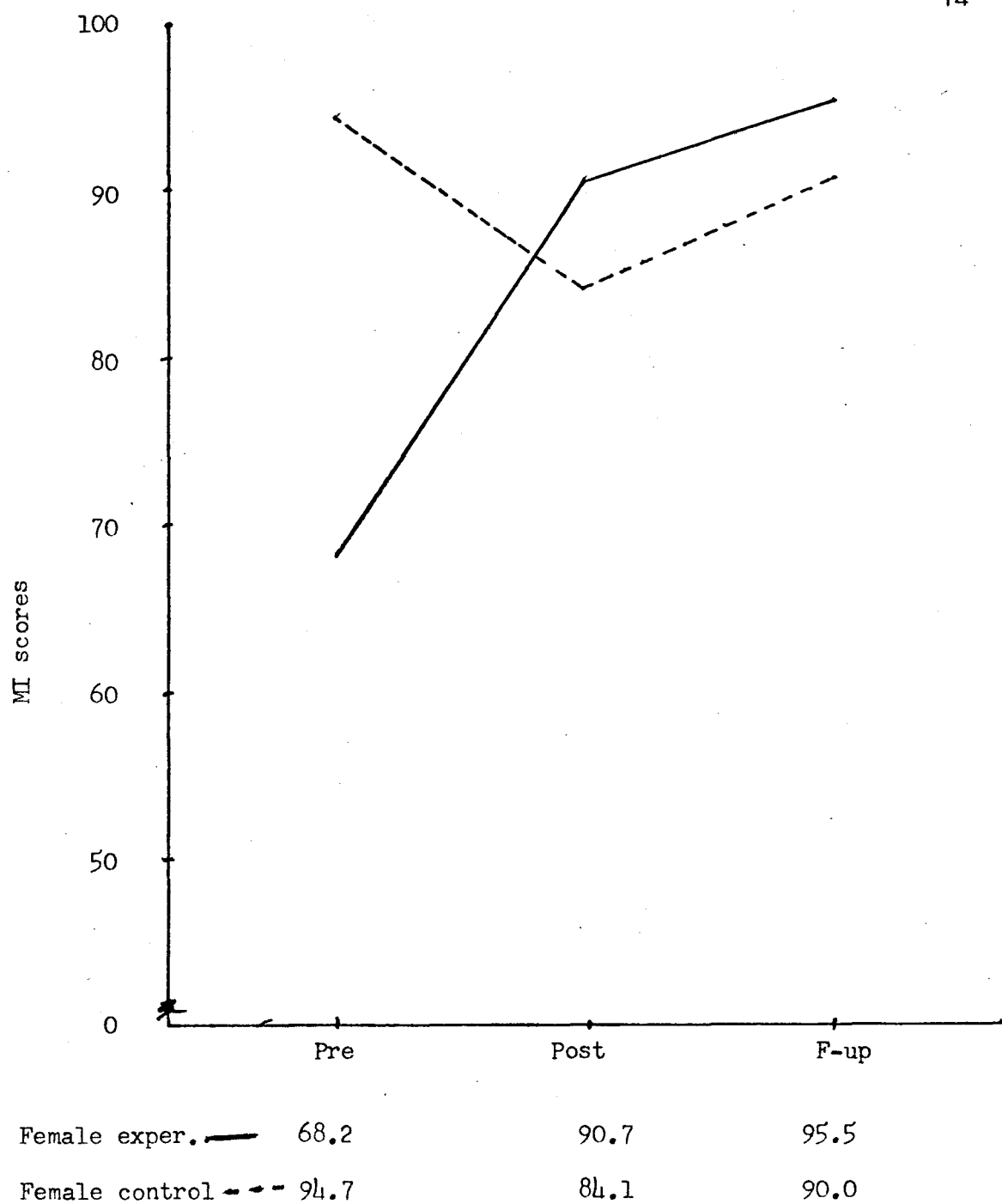


Figure 6. Median MI scores of female experimental and female control subjects.

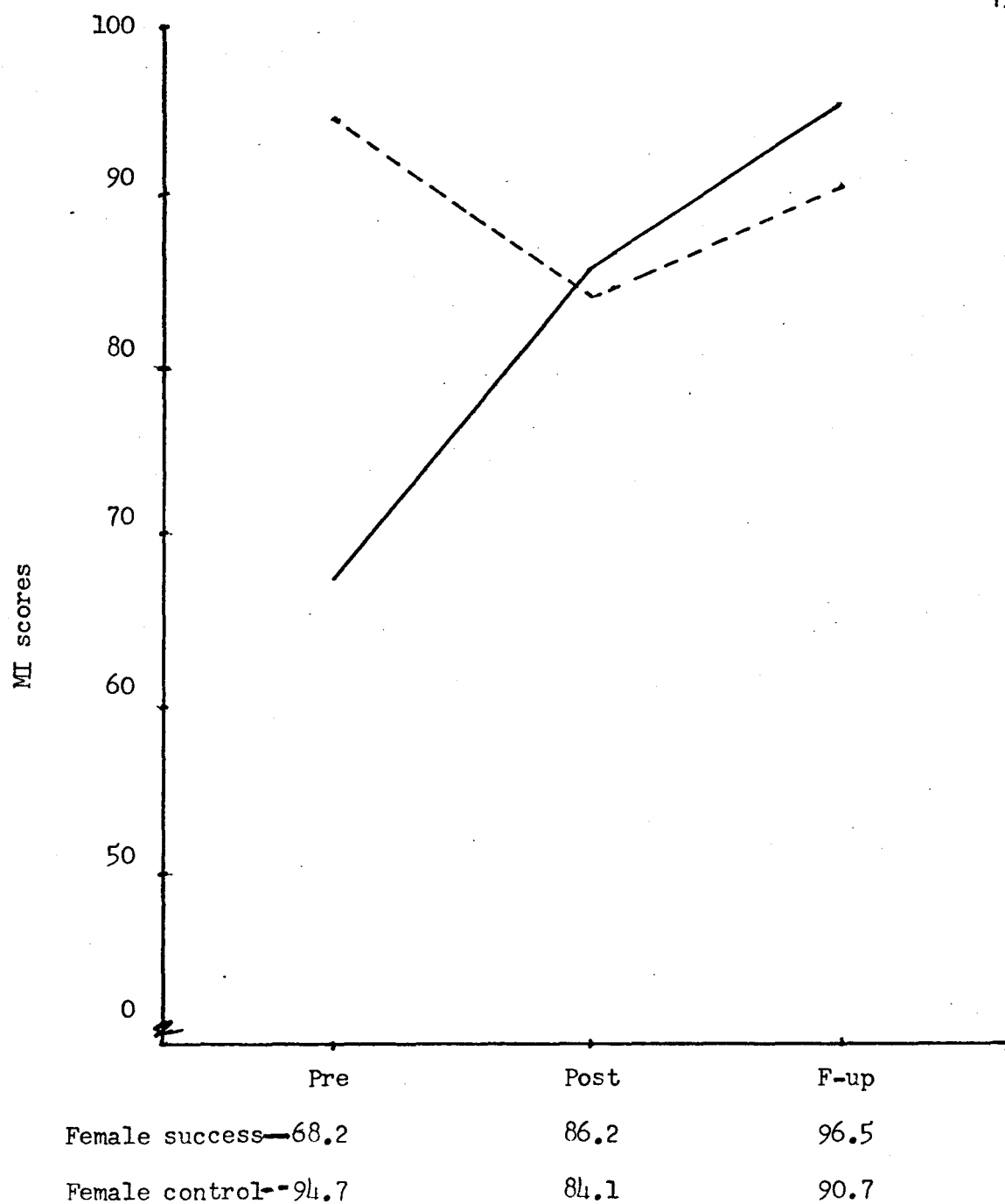


Figure 7. Median MI scores of female success experimental and female control subjects.

than the males, ($z = 2.21$, $p = .03$, two-tailed test. See Table 4.) This can be understood if it is recalled that the male success Ss increased their MI scores significantly at post-therapy (Wilcoxon $z = 1.68$, $p < .05$), but decreased them from post-therapy to follow-up. The female success subgroup, on the other hand, increased their MI scores from pre-therapy to post-therapy and from pre-therapy to follow-up. (Wilcoxon z 's were: 2.63 and 2.52, $p < .01$). From post-therapy to follow-up their MI increase had only a probability of .08 ($z = 1.39$). See Table 3, Figure 8.

Figure 8

Finally, the male experimental Ss were divided into male success and male failure subgroups and were compared with each other at pre-therapy, post-therapy and follow-up by the Mann-Whitney U Test. They are not statistically different at pre-therapy. At post-therapy the MI scores of the male success subgroup are higher than those of the male failure Ss, and their difference from the male failure Ss is significant at the .05 level in a two-tailed test ($z = 2.08$, $p = .04$). This is attributed to therapy. At follow-up their difference is not statistically significant due to the collapse of the MI scores of the male success group after therapy. (See Table 4, Figure 9.)

Figure 9

Looking at Figures 8 and 9, the investigator was impressed by the sensitivity and reliability of the Story Sequence Analysis as an instrument for the study of therapy. It can be conjectured what happened to the male experimental group. The failure cases decreased their MI constantly;

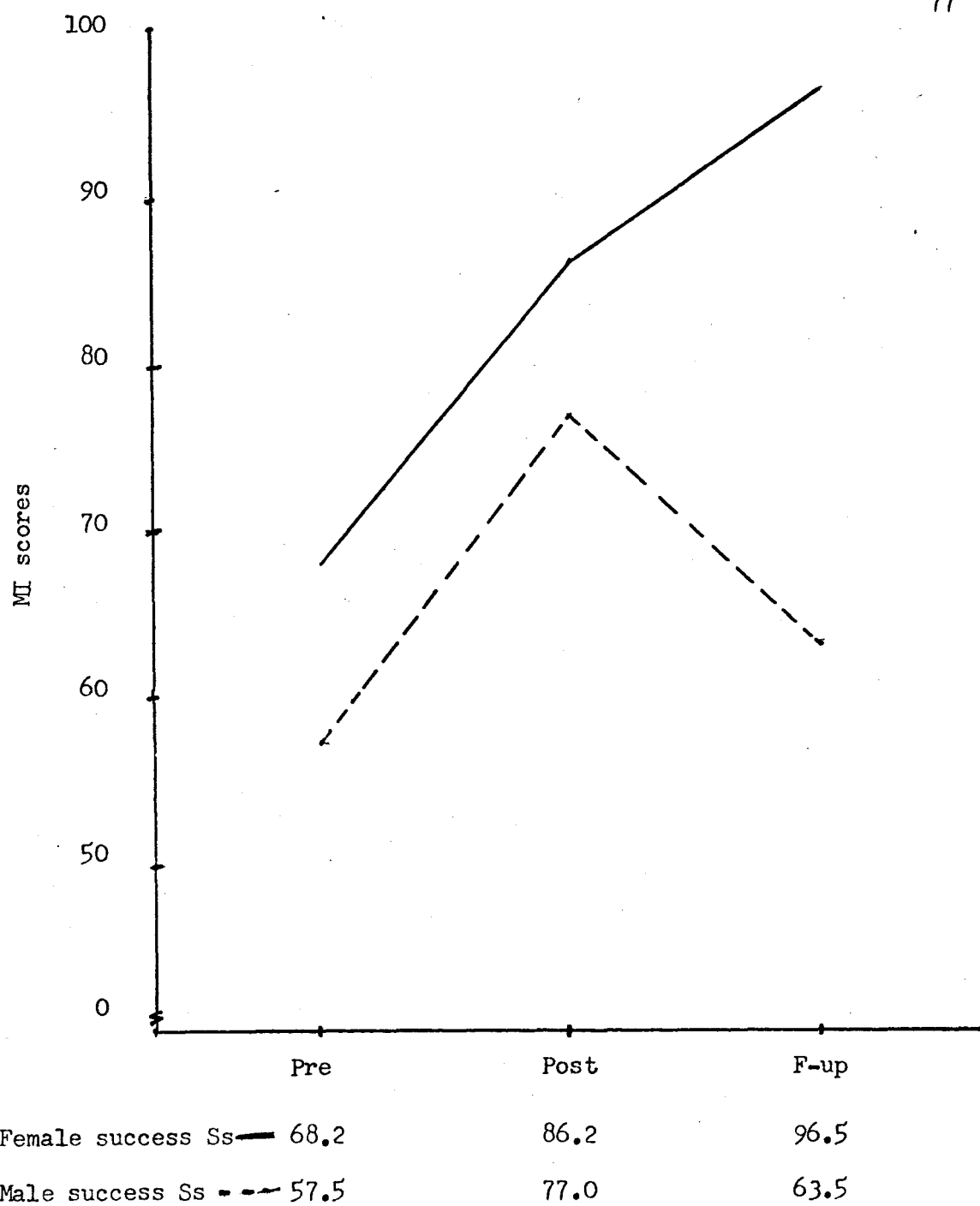
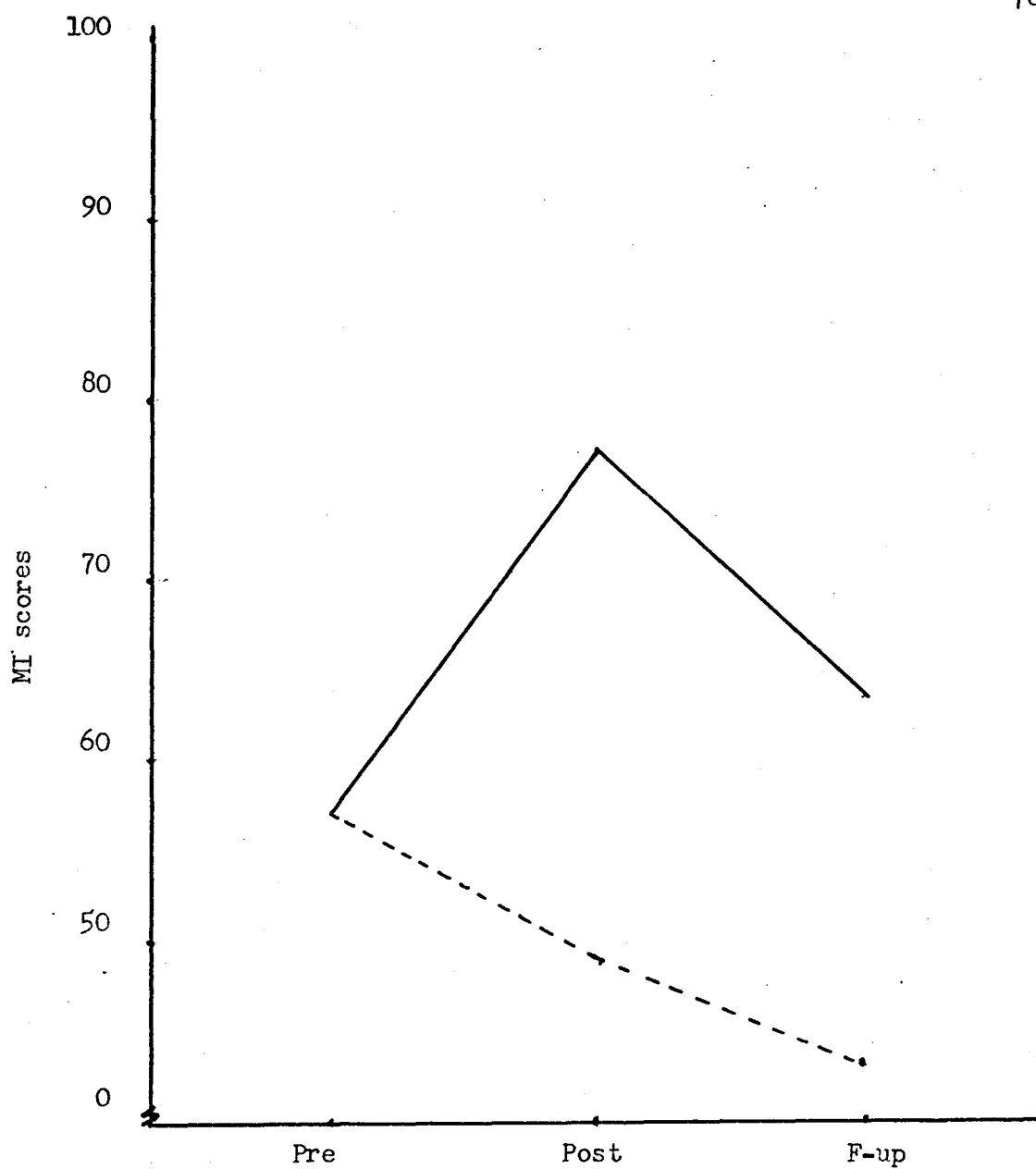


Figure 8. Median MI scores of female success and male success experimental Ss.



Male success Ss — 57.0

77.0

63.5

Male failure Ss --- 57.0

49.0

43.0

Figure 9. Median MI scores of male success and failure experimental Ss.

the male success group increased their MI scores during therapy, their progress being beautifully parallel to that of the female success Ss. After therapy, the female success Ss continued their way up giving some support to Rogers' contention that therapy starts a process that continues after the therapy sessions are terminated. The male success group, on the other hand, collapsed after therapy. The reason for this can not be found in this study. However, since all the therapists were males, it is possible that they could not establish a meaningful relationship with the failure Ss, and that the positive relationship they succeeded to develop with the male success subjects during therapy was not strong enough to hold up after therapy. It could also be speculated that this heterosexual relationship between the success female clients and the male therapists was so strong and meaningful that it helped the female clients to carry on and upwards after therapy. These possibilities should be studied by an adequate research design. It suffices to state here that the reliability and sensitivity of the SSA should be of great help in such a study.

The findings from the multiple comparisons made between the experimental and control subjects and between the success and failure cases across the sex variable can be summarized in the following short statements:

1. The MI scores of the female experimental subjects are higher than those of the male experimental Ss. The difference is significant at post-therapy and follow-up.
2. The MI scores of the female control Ss are also higher than those of the male control Ss. This difference is significant at pre-therapy.

3. The female and male subjects rated as success by their therapists are not statistically different at pre-therapy and post-therapy though the MI scores of the female Ss are higher. At follow-up, however, the difference is significant at the .05 level as a result of the continuous progress of the female success Ss after therapy and the decrease of the MI scores of the male Ss.

4. The female Ss of the experimental group differed significantly from the female Ss of the control group at pre-therapy ($p \leq .05$) but not at post-therapy and follow-up. The female experimental Ss were significantly lower at pre-therapy but were higher than the female control Ss at post-therapy and follow-up.

5. The female experimental Ss significantly increased their MI scores from pre-therapy to post-therapy and from pre-therapy to follow-up. This is attributed to therapy.

6. The female success experimental Ss differed significantly from the female control Ss at pre-therapy ($p \leq .05$) in that they had significantly lower MI's. At post-therapy and follow-up, their MI's were higher than those of the female controls.

7. The female success experimental Ss significantly increased their MI scores from pre-therapy to post-therapy ($p < .01$), from pre-therapy to follow-up ($p = .01$). This increase is attributed to therapy.

8. The female control Ss had higher MI scores than the female experimental Ss at pre-therapy ($p \leq .05$), but they decreased their MI scores continuously from pre-therapy to post-therapy and from pre-therapy to follow-up. The decrease at post-therapy is significant at the .05 level.

9. The male experimental Ss did not differ significantly from the male control Ss at any of the three points of comparison. Both groups showed random changes from test to test.

10. The male success Ss increased their MI scores from pre-therapy to post-therapy at the .05 level of significance. This is attributed to therapy.

11. The male success subjects did not differ significantly from the male failure subjects at pre-therapy but they differed significantly at post-therapy ($p = .04$).

12. While the male success group increased their MI scores from pre-therapy to post-therapy at the .05 level of significance, the male failure subjects significantly decreased their MI scores from pre-therapy to post-therapy ($p = .05$).

It can be concluded, therefore, that the male and female Ss who were rated as success cases by their therapists significantly improved their MI scores during therapy. That this positive change was due to therapy and not to sex is further evidenced by the fact that the female control Ss and the failure experimental Ss decreased their MI scores significantly during the period of therapy. The only exception would be the two female experimental Ss (Nos. 18 and 21, Table 1) rated as failure by their therapists. They also increased their MI scores from pre-therapy to post-therapy and could be considered success cases by the SSA. The investigator checked their scores on the other psychological tests they took together with the TAT. Both increased also their Self-Ideal Self correlations and their E-M ratings from pre-therapy to post-therapy. Possibly the counselor's

rating was too stringent in these two cases. The male control subjects showed only random changes.

After this detailed discussion of the sex variable and the MI scores of the experimental and control Ss, it can be restated more clearly that the four main hypotheses of the study are confirmed.

Similar investigations could be made along the other actuarial variables used to match the experimental and control Ss, but because of the limited number of subjects in this study it would be very difficult to partial out the influence of each variable and leave a substantial number of subjects in each subgroup to work on and obtain reliable results.

CHAPTER VI

SUMMARY, CONCLUSIONS AND SUGGESTIONS FOR FUTURE RESEARCH

I. SUMMARY

A. Purpose of the Study

The present study was designed to test the hypothesis that successful therapy results in more positive motivational attitudes as measured by the Story Sequence Analysis (SSA). It also purported to test the validity of the SSA as a measure of the outcome of therapy by comparing the scores of the Motivation Index (MI) with the scores obtained by the subjects in other psychological tests used in the evaluation of client-centered therapy.

B. Variables

The independent variable of the study was client-centered therapy, and the basic criterion of the success in therapy was the rating done by the therapist at the end of therapy. The dependent variable was motivation. It was operationally defined as convictions and attitudes revealed by the imports of the TAT stories. It was measured by Arnold's method of Story Sequence Analysis and was scored along four categories: 1) Achievement, success and their opposites; 2) Right and wrong intentions and actions; 3) Human relations; and 4) Reactions to adversity (Arnold, 1962)

C. Subjects

The experimental Ss are thirty out-patients at the Counseling and Psychotherapy Research Center of the University of Chicago. Eleven are female and nineteen are male, with an age range of 20 to 40 years and a median of 26.8 years. Sixteen are graduate or undergraduate students at the University of Chicago and fourteen are in different professions and

occupations in the community. Twenty-six subjects, matched with the experimentals in sex, age, socio-economic status and education, served as controls. They took the TAT and a battery of psychological tests at approximately the same time interval as the experimentals but did not take therapy.

D. Research Material

A 20-card TAT was administered individually to the experimental and control Ss at pre-therapy, post-therapy and follow-up, six to twelve months after the termination of therapy. The TAT stories were given orally and recorded on tape. They were later transcribed by the clerical staff of the University of Chicago Counseling Center.

All the TAT records had been coded by other investigators. They were first analyzed blindly by this investigator. The correlation between the first eleven and the total of twenty stories in 100 records scored by the investigator was .956. Fifteen randomly selected TAT protocols were analyzed by an independent rater. His correlations of the first eleven stories with the total of all twenty stories was .97. Consequently, only the score derived from the first eleven stories of each record was used to test the hypothesis of this study.

The interjudge reliability of the investigator was tested by comparing his scores of the first analysis with the score of forty TAT records which were rescored six to twelve months later. The correlation was .75. His interjudge reliability was calculated by comparing his scores with the scores of two independent raters who analyzed and scored blindly thirty-six TAT records selected by random numbers from the experimental and control groups. The interjudge reliability of the investigator and rater B was

.84; of the investigator and rater C was .70, and of the two independent raters, .66. The interjudge reliability of the three raters calculated by the method suggested by Guilford was .92 (Guilford, 1955, p. 35.)

E. Hypotheses of the study

Hypothesis 1: The experimental Ss who received therapy for at least six session will have significantly higher MI scores at post-therapy and follow-up than at pre-therapy.

The increase of the MI scores actually found was tested by the Wilcoxon Signed-Ranks Test. At post-therapy it was significant at the .05 level, and was presumably attributable to therapy.

Hypothesis 2: The control Ss who took the TAT at approximately equal time intervals as the experimental Ss but had no therapy will not show any significant increase in their MI scores at post-therapy and follow-up.

Instead of an increase, they showed a significant decrease in their MI scores at post-therapy, calculated by the Wilcoxon Test.

Hypothesis 3: Experimental Ss rated as success by their therapists will have significantly higher MI scores at post-therapy and follow-up than at pre-therapy.

The hypothesis was tested by the Wilcoxon Test. It was confirmed at the .001 level at post-therapy and at the .05 level at follow-up.

Hypothesis 4: Experimental Ss rated as failure cases by their therapists will not show significantly higher MI scores at post-therapy and follow-up than at pre-therapy.

Instead of an increase, they had a decrease in their MI scores which was significant at the .05 level. The hypothesis was confirmed.

Hypothesis 5: Success experimental Ss will have higher MI scores than failure cases at pre-therapy.

The difference was tested by the Mann-Whitney U Test. It was not statistically significant. This hypothesis was not confirmed.

F. Comparison of the SSA with other psychological tests

The MI scores of the experimental and control Ss were compared with the scores they received on the Self-Ideal Self Q Sort, the Emotional Maturity Scale, the Counselors' Rating, the Personal Integration Rating and the TAT analyzed and scored by a clinical method. The MI scores correlated significantly with the TAT scores, the Personal Integration Rating and the Counselors' Ratings. No significant correlation was found with the Self-Ideal Self Q Sort and the E-M scale.

II. CONCLUSIONS

Only the experimental Ss rated as success therapy cases--male and female--had a significant increase in their MI scores at post-therapy.

The experimental Ss rated as failure cases by their therapists had a significant decrease in their scores at post-therapy.

The control Ss showed also a significant decrease in their MI scores at post-therapy.

Though applied to therapy cases for the first time, the SSA has proved to be an adequate, reliable and valid instrument for the study of the outcome and the process of therapy. It has shown an important heuristic value in the study of the opposite effects therapy seems to have on success and failure cases.

III. SUGGESTIONS FOR FUTURE RESEARCH

Since the limited number of Ss in this study makes it impossible to partial out the effect of age, sex and education on the MI change in success and failure cases, it would be desirable to establish norms for men and women, as well as for different age and educational groups, on the basis of a larger unselected sample. Although it seems reasonable to expect an increase in MI with increasing age in a normal group, the small number of cases in the control group makes it impossible to say so with any certainty. Whether women and non-students in the general population have higher MI's than men and students would be another question that would deserve a larger normative survey.

This investigator was greatly and happily surprised to evidence the reliability and validity of the results obtained by the strict application of the SSA to therapy cases. The SSA is an outstanding instrument for future research on the outcome and process of therapy. May he express his hope that future investigators find ways and means to extend the use of the SSA to therapy cases and improve on the method itself where there is room for improvement.

BIBLIOGRAPHY

- Abt, L. & Ballak, L., Projective Psychology. Clinical Approaches to the Total Personality. New York. Alfred Knoff, 1950.
- Arnold, Magda B., "A Demonstration Analysis of the Thematic Apperception Test in a Clinical Setting". Journal of Abnormal and Social Psychology. 1949, 44: 97-111.
- Arnold, Magda B., "Analytic Technique". In E. C. Sneidman. Thematic Test Analysis. New York. Grune & Stratton, 1951, pp. 31-39.
- Arnold, M. B. & Casson, J. A., The Human Person. New York. Ronald Press, 1954.
- Arnold, Magda B., Emotion and Personality, 2 vols. New York. Columbia University Press, 1960.
- Atkinson, J. W., "Exploration Using Imaginative Thought to Assess the Strength of Human Motives". In M. R. Jones (ed) Nebraska Symposium on Motivation Lincoln. University of Nebraska Press, 1954, pp. 56-112.
- Atkinson, John W. (ed), Motives in Fantasy, Action and Society. New York. Van Nostrand Col, 1958.
- Atkinson, J. W., "Thematic Apperceptive Measurement of Motives with the Context of a Theory of Motivation". In J. W. Atkinson (ed) Motives in Fantasy, Action and Society. New York. Van Nostrand, 1958, pp. 519-616.
- Atkinson, J. W., "Discussion of Dr. Lazarus Paper 'A Substitutive-Defensive Conception of Apperceptive Fantasy'". In J. Kogan & G. Lesser, Contemporary Issues in Thematic Apperceptive Methods. Springfield, Ill. 1961, pp. 72-82.
- Barrington, Byron L., "The Differential Effectiveness of Therapy as Measured by the Thematic Apperception Test". In C. Rogers et al. The Therapeutic Relationship and its Impact. Madison. The University of Wisconsin Press, 1967, pp. 337-352.
- Barron, F. & Leary T. F., "Changes in Psychoneurotic Patients with and without Psychotherapy". Journal of Consulting Psychology. 1955, 19:239-245.
- Bellak, Leopold, "The Thematic Apperception Test in Clinical Use". In L. Abt & L. Bellak, Projective Psychology. New York. Knoff, 1950, pp. 185-229.

- Bettelheim, Bruno, "Self-Interpretation of Fantasy: the Thematic Apperception Test as an Educational and Therapeutic Device". American Journal of Orthopsychiatry. 1947, 17:80-100.
- Bimey, R. & Teevan, R., Measuring Human Motivation. New York. Van Nostrand, 1962.
- Borgatta, E. F. & Cottrell, L. S., "Control Group Experimentation in Psychotherapy". Psychiatry. 1959, 22:97-100.
- Burkard, Mary I., "Characteristic Differences Determined by TAT Sequence Analysis between Teachers Rated by their Pupils at the Extremes in Teaching Efficiency". Ph.D. Diss. Loyola University, Chicago, 1958.
- Brown, Joseph E., "Personality Dynamics of High and Low Academic Achievers in High School: A Modified Thematic Apperception Test Reliability Study". Master's Thesis. St. Louis University, St. Louis, 1953.
- Butler, J. & Haigh, G., "Changes in the Relation between Self-Concepts and Ideal Concepts upon Client-centered Counseling". In C. Rogers & R. Dymond, Psychotherapy and Personality Change. Chicago. University of Chicago Press, 1954, pp. 55-75.
- Calvin, A. D., "Some Misuses of the Experimental Method in Evaluating the Effect of Client-centered Counseling". Journal of Consulting Psychology, 1954, 1:249-251.
- Cartwright, Desmond, "Success in Psychotherapy as a Function of Certain Actuarial Variables". Journal of Consulting Psychology, 1955, 5:357-363.
- Cartwright, D., "A Rapid Nonparametric Estimate of Multi-judge Reliability". Psychometrika, 1956, 21:17-29.
- Cartwright, D. S., "Annotated Bibliography of Research and Theory Construction in Client-centered Therapy". Journal of Counseling Psychology, 1957, 4:82-100.
- Cartwright, D., Kirtner, W. & Fiske, D., "Method Factors in Changes Associated with Psychotherapy". Journal of Abnormal Social Psychology, 1963, 66:164-175.
- Cartwright, Rosalind, & Vogel, J., "A Comparison of Changes in Psycho-neurotic Patients during Matched Periods of Therapy and No-therapy". Journal of Consulting Psychology, 1960, 24:121-127.
- Cross, Herbert J., "The Outcome of Psychotherapy: a Selected Analysis of Research Findings". Journal of Consulting Psychology, 1964, 28:413-417.

- Dana, Richard, "Cross Validation of Objective TAT Scoring". Journal of Consulting Techniques. 1956, 20:33-36.
- Dana, Richard H., "An Application of Objective TAT Scoring". Journal of Projective Techniques. 1956, 20:159-163.
- Deabler, Herdis L., "The Psychotherapeutic Use of the Thematic Apperception Test". Journal of Clinical Psychology, 1947, 3:246-252.
- Dinoff, M., Richard H., Salzbert, H., & Sipprelle, C., "An Experimental Analogue of Three Psychotherapeutic Approaches". Journal of Clinical Psychology, 1960, 16:70-73.
- Dymond, Rosalind, "Adjustment Changes over Therapy from Thematic Apperception Test Ratings". In C. Rogers & R. Dymond, Psychotherapy and Personality Change. Chicago: University of Chicago Press, 1954, pp. 109-120.
- Dollard, J. & Auld, F., Scoring Human Motives: A Manual. New Haven: Yale University Press, 1959.
- Dore, Patricia D., "The Characteristic Pattern of Attitudes of Petty Offenders and Men on Relief as Revealed by TAT Story Sequence Analysis". Ph.D. Diss. Loyola University, Chicago, 1968.
- Dulin, John, S.J., "Prediction of Academic Achievement in College". Ph.D. Diss. Loyola University, Chicago, 1968.
- Edwards, A. & Cronbach, L., "Experimental Design for Research in Psychotherapy". Journal of Clinical Psychology, 1952, 8:51-59.
- Ehrlich, D., & Wiener, D., "The Measurement of Values in Psychotherapeutic Setting". Journal of General Psychology, 1961, 69:359-372.
- Eysenck, H. J., "The Effects of Psychotherapy: An Evaluation". Journal of Consulting Psychology. 1952, 16:319-323.
- Fairweathers, G. & Simon R., "A Further Follow-up Comparison of Psychotherapeutic Programs". Journal of Consulting Psychology, 1963, 27:186.
- Fairweathers, G. W. et al., "Relative Effectiveness Comparison of Four Programs for Three Different Patients Groups". Psychological Monographs, 1960, No. 492.
- Fields, Rona, "A Personality Description of the Unwed Pregnant Adolescent Using Story Sequence Analysis". M.A. Thesis. Loyola University, Chicago, 1965.

- Fiske, D., Cartwright, D. & Kirtner, W., "Are Psychotherapy Changes Predictable?" Journal of Abnormal and Social Psychology, 1964, 69:418-426.
- Frank, Jerome, "Problem of Controls in Psychotherapy as Exemplified by the Psychotherapy Research Project of the Phipps Psychiatric Clinic". In E.A. Rubinstein & M. Parloff (eds), Research in Psychotherapy, Vol. 1 Washington, D.C. APA. 1958, pp. 10-26.
- Garvin, Joseph, "A Thematic Apperception Test Study of Non-intellective Factors Related to Academic Success on the College Level". Ph.D. Diss., Loyola University, Chicago, 1958.
- Green, B.F., "Attitude Measurement". In G. Lindzey (ed), Handbook of Social Psychology. Cambridge: Addison Wesley, 1954, Vol. 1, pp. 335-369.
- Grummon, D. & John, Eve, "Changes over Client-centered Therapy Evaluated on Psychoanalytically Based Thematic Apperception Test Scales". In C. Rogers & R. Dymond (eds), Psychotherapy and Personality Change. Chicago: University of Chicago Press, 1954, pp. 121-144.
- Guilford, J.P., Psychometric Methods. New York: McGraw-Hill, 1954.
- Hathaway, S. R., "Some Considerations Relative to Nondirective Counseling as Therapy". Journal of Clinical Psychology, 1948, 4:226-231.
- Henry, W. & Shlien, J., "Affective Complexity and Psychotherapy: Some Comparisons of Time-Limited and Unlimited Treatment". Journal of Projective Techniques. 1958, 22:153-162.
- Hobbs, N., "Sources of Gain in Psychotherapy". American Psychologist, 1962, 17:741-747.
- Holt, R.R., "The Nature of TAT Stories as Cognitive Products: a Psychoanalytic Approach". In J. Kagan & G. Lesser (eds), Contemporary Issues in Thematic Apperceptive Methods. Springfield: Thomas, 1960, pp. 3-51.
- Jensen, A.R., The Reliability of Projective Techniques: Methodology. Amsterdam: North Holland Publishing Company. 1959.
- Kogan, J., & G. Lesser (Eds) Contemporary Issues in Thematic Apperceptive Methods. Springfield: Thomas. 1961.
- Korner, I.M., Research on Processes and Outcomes of Psychotherapy. Lawrence: University of Kansas, 1957.

- Lazarus, R.S., "A Substitutive-Defensive Conception of Apperceptive Fantasy". In J. Kagan & G. Lesser (eds), Contemporary Issues in Thematic Apperceptive Methods. Springfield: Thomas, 1961, pp. 51-71.
- Lasaga y Travieso, J. & Martinez Arango, C., "Some Suggestions Concerning the Administration and Interpretation of the Thematic Apperception Test". Journal of Psychology, 1946, 22:117-163.
- Lasaga y Travieso, J. & Martinez Arango, C., "Four Detailed Examples of How Mental Conflicts of Psychoneurotic and Psychotic Patients may be Discovered by Means of the Thematic Apperception Test". Journal of Psychology, 1948, 26:299-345.
- Little, K. & Shneidman, E., "The Validity of Thematic Projective Technique Interpretation". Journal of Personality, 1955, 23:285-294.
- Luborsky, L., "Self-interpretation of the TAT as a Clinical Technique". Journal of Projective Techniques, 1953, 17:217-223.
- McCandlish, L.A., "An Investigation of a New Method of TAT Analysis". Ph.D. Diss., Loyola University, Chicago, 1958.
- McClelland, D., Atkinson, J., Clark, R. & Lowell, E., The Achievement Motive. New York: Appleton-Century, 1953.
- Meyers, J. & Auld, F., "Some Variables Related to Outcome of Psychotherapy". Journal of Clinical Psychology. 1955, 11:51-54.
- Meyers, M. & Tolman, Ruth, "Correspondence between Attitudes and Images of Parental Figures in TAT Stories and in Therapeutic Interviews". Journal of Consulting Psychology, 1955, 19:79-82.
- Morgan, D. & Murray, H., "A Method for Investigating Fantasies: The Thematic Apperception Test". In Birney, R. & Teevan, R., Measuring Human Motivation. New York: Van Nostrand, 1962, pp. 62-86.
- Morton, R. B., "An Experiment in Brief Psychotherapy". Psychological Monographs, 1955, (Whole No. 386).
- Murray, H.A., Thematic Apperception Test Manual. Cambridge: Harvard University Press, 1943.
- Murstein, Bernard, Theory and Research in Projective Techniques. New York: Wiley and Sons, 1963.
- Najjar, B., "Difference in Motivational Patterns of Inadequate and Inept Psychopathic Deviant Prison Inmates as Revealed by TAT Story Sequence Analysis". Ph.D. Diss. Loyola University, Chicago, 1967.

- Petrauskas, F., "A TAT and Picture-Frustration Study of Naval Offenders". Ph.D. Diss., Loyola University, Chicago, 1958.
- Quinn, T.L., "Differences in Motivational Patterns of College Student Brothers as Revealed in the TAT, the Ratings of their Peers and the Ratings of their Superiors: A Validation Study". Ph.D. Diss., Loyola University, Chicago, 1961.
- Reitman, W.R., & Atkinson, J.W., "Some Methodological Problems in the Use of Thematic Apperceptive Measures of Human Motives". In J.W. Atkinson (ed) Motives in Fantasy, Action and Society. Princeton: Van Nostrand, 1958, pp. 664-683.
- Ricks, D., Umbarger, C. & Mach, R., "A Measure of Increased Temporal Perceptive in Successfully Treated Adolescent Delinquent Boys". Journal of Abnormal Social Psychology, 1964, 69:685-689.
- Rogers, C., Client-centered Therapy, New York: Houghton Mifflin Co., 1951.
- Rogers, Carl, "Changes in the Maturation of Behavior as Related to Therapy". In C. Rogers & R. Dymond, Psychotherapy and Personality Change. Chicago: University of Chicago Press, 1954, pp. 215-237.
- Rogers, C. & Dymond, R. (eds), Psychotherapy and Personality Change. Chicago: University of Chicago Press, 1954.
- Rogers, Carl et al., The Therapeutic Relationship and its Impact. Madison: University of Wisconsin Press, 1954.
- Sargen, Helen, "Intra-Psychic Change: Methodological Problems in Psychotherapy Research". Psychiatry, 1961, 24:93-108.
- Seeman, J., "Counselor Judgments of Therapeutic Process and Outcome". In C. Rogers & R. Dymond (eds), Psychotherapy and Personality Change. Chicago: University of Chicago Press, 1954, pp. 99-108.
- Scheidman, Edwin et al. (eds), Thematic Test Analysis. New York: Grune & Stratton, 1951.
- Snider, L.B., "A Research Method Validating Self-Determination". In M. Arnold & J. Gasson (eds), The Human Person. New York: Ronald, 1954, pp. 222-263.
- Steggert, F., "An Analysis of Some Personal and Executive Characteristics of Participants in a University Program of Executive Development for Federal Personnel". Ph.D. Diss. Loyola University, Chicago, 1961.

- Stieper, D.R. & Wienen, D.N., Dimensions of Psychotherapy. An Experimental and Clinical Approach. Chicago: Aldine Publishing Co., 1965.
- Strupp, H.H., "The Outcome Problem in Psychotherapy Revisited". Psychotherapy: Theory, Research and Practice, 1963, 1:1-13.
- Ullmann, L.P., & McFarland, R.L., "Productivity as a Variable in TAT Protocols: A Methodological Study". Journal of Projective Techniques. 1957, 21:80-87.
- Vassiliou, V. "Motivational Patterns of Two Clinical Groups as Revealed by TAT Sequence Analysis". Ph.D., Diss. Loyola University of Chicago, 1962.
- Volsky, T., Magoon, T., Norman, W. & Hoyt, W., The Outcome of Counseling and Psychotherapy. Theory and Research. Minneapolis: University of Minnesota Press, 1965.
- Wallach, M.S. & Strupp, H.H., "Psychotherapists' Clinical Judgments and Attitudes Toward Patients". Journal of Consulting Psychology, 1960, 24:316-323.
- Watterson, D.J., "Problems in the Evaluation of Psychotherapy". Bulletin of Menninger Clinic, 1954, 18:232-241.
- Williams, J.E., "Changes in Self and Other Perceptions Following Brief Educational Vocational Counseling". Journal of Counseling Psychology, 1962, 9:18-28.
- Wyatt, F., "The Scoring and Analysis of the Thematic Apperception Test". Journal of Psychology, 1947, 24:319-330.
- Wyatt, F. & Veroff, J.B., "Thematic Apperception and Fantasy Tests". Progress of Clinical Psychology, 1956, 2:32-57.
- Wyatt, F.A., "A Principle for the Interpretation of Fantasy". Journal of Projective Techniques, 1958, 22:229-245.
- Zax, M. & Klein, A., "Measurement of Personality and Behavior Changes Following Psychotherapy". Psychological Bulletin, 1960, 57:435-448.
- Zubin, J., "Evaluation of Therapeutic Outcome in Mental Disorders". Journal of Nervous Mental Diseases, 1953, 117:95-111.

APPROVAL SHEET

The dissertation submitted by Augustine Ramirez, O.F.M. has been read and approved by members of the Department of Psychology.

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

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

1/21/1970

Date

Magda B. Arnold

Signature of Advisor