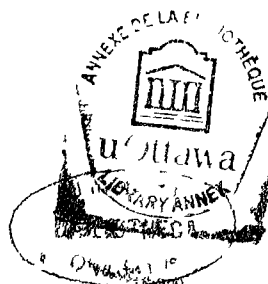


**SELF-REFERENT VERSUS OTHER-REFERENT INCOMPLETE
SENTENCES IN PROJECTION**

by **Arthur C. Keating**

**Thesis presented to the School of Psychology and
Education of the University of Ottawa as partial
fulfillment of the requirements for the degree
of Master of Arts in Psychology.**

Ottawa, Ontario, 1960



UMI Number EC55202

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform EC55202
Copyright 2011 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

ACKNOWLEDGEMENTS

This thesis was prepared under the guidance of Mr. Gilles Chagnon, M.Ps., of the School of Psychology and Education, University of Ottawa.

The writer is also indebted to Dr. Lawrence Dayhaw of the same School for his invaluable advice, and to his former colleagues in the Royal Canadian Air Force who arranged for the test administration.

CURRICULUM STUDIORUM

Arthur C. Keating was born February 25, 1926 in Hove, England. He received the degree of Bachelor of Science in Education in September 1954, and the degree of Bachelor of Science, Honours in Psychology, in May 1958, from the University of Ottawa.

TABLE OF CONTENTS

Chapter	page
INTRODUCTION	vii
I.-THE RATIONALE UNDERLYING SENTENCE COMPLETION TESTS	1
1. The Concept of Projection	1
2. Personality Theory and Projective Techniques	9
3. Sentence Completion Tests as Projective Techniques	11
II.-GENERAL STUDIES OF THE SENTENCE COMPLETION METHOD	16
1. Literature to 1947	16
2. Literature from 1948 to 1957	33
III.-REVIEW OF LITERATURE RELATED TO TYPES OF ITEMS	52
1. General	52
2. Supporting the Use of Self-Referent Items	57
3. Supporting the Use of Other-Referent Items	62
4. Supporting the Use of Both Types of Items	75
IV.-THE PRESENT STUDY	80
1. Statement of the Problem	80
2. The Sample Groups	87
3. Procedure	92
4. Statistical Treatment of the Data	94
V.-FINDINGS AND DISCUSSION	97
1. Interscorer Reliability	97
2. Test Reliability	98
3. Null Hypotheses One and Two	98
4. Null Hypotheses Three and Four	103
5. General Discussion	104
SUMMARY AND CONCLUSIONS	109
BIBLIOGRAPHY	112

TABLE OF CONTENTS

v

page

Appendix

1.	<u>EXPERIMENTAL SENTENCE COMPLETIONS, FORM M</u>	118
2.	<u>RATING RECORD SHEET, SENTENCE COMPLETIONS -</u> <u>OTHER-REFERENT FORM</u>	124
3.	<u>RATING RECORD SHEET, SENTENCE COMPLETIONS -</u> <u>SELF-REFERENT FORM</u>	126
4.	<u>ABSTRACT OF SELF-REFERENT VERSUS OTHER-REFERENT</u> <u>INCOMPLETE SENTENCES IN PROJECTION</u>	128

LIST OF TABLES

Table	page
I.- Means and Differences between Means of Negative Other-Referent and Self-Referent Scores for the Adjusted Group	100
II.- Comparison of the Variances within the Adjusted Group on the Self-Referent and Other-Referent Items	100
III.- Means and Differences between Means of Negative Other-Referent and Self-Referent Scores for the Maladjusted Group	102
IV.- Comparison of the Variances within the Maladjusted Group on the Self-Referent and Other-Referent Items	102
V.- Means and Differences between Means of the Maladjusted and Adjusted Groups on the <u>Self-Referent Form of the Experimental</u> <u>Sentence Completion Test</u>	105
VI.- Means and Differences between Means of the Maladjusted and Adjusted Groups on the <u>Other-Referent Form of the Experimental</u> <u>Sentence Completion Test.</u>	105

INTRODUCTION

An assumption, widely held among psychologists is that the sentence completions of an individual reveal latent needs, sentiments, attitudes and aspirations which that person would be unwilling or unable to recognize or express in response to direct questioning. From this assumption, it has been reasoned that sentence completion tests are projective techniques of personality assessment.

Sentence completion items can be classified as personal or impersonal according to whether the stems contain either a person's name or a personal pronoun, or are neutral and make no reference to persons, real or fictitious. Two types of personal items may be used: self-referent stems, which contain a first personal pronoun, and other-referent stems, that utilize a third personal pronoun, or a proper name. A study by Joseph Sacks of these two types of personal items resulted in a report being published in 1949. In this report, it was noted that the self-referent type of item yielded the greatest degree of similarity with psychologists' and psychiatrists' impressions of a group of patients at a Mental Hygiene Clinic.

Since that time, various authors have interpreted this report as meaning that self-referent items are superior

to other-referent items; that the former are more projective than the latter. However, there have been others who assumed that a subject revealed more about himself when completing the other-referent stems than when completing self-referent items. Needless to say, comparatively little research has been carried out by either camp.

It would appear that if the evaluations based on the responses to incomplete sentences are to have any validity, it is essential to investigate the assumptions underlying this instrument.

This investigator has undertaken to report on the projective nature of self-referent and other-referent sentence completion items, using not a subjective, introspective judgment of their natures, but objective, quantitative methods. The instrument employed is a modification of that utilized by Sacks in his experiment.

Chapter I of this report will present a discussion of the concept of projection in its application to personality assessment, and in particular to the sentence completion method as found in the literature to date.

Chapter II will review the use of the technique to 1957, giving special emphasis to the validity and the reliability of the various forms of the test.

Chapter III will be devoted to reviewing those studies relating to the use of different types of items, in particular to the self-referent and other-referent stems.

The rationale underlying the present investigation, a statement of the problem and the hypotheses which evolved from it, will appear in Chapter IV. Besides this, the method and criteria used to select the experimental groups will be described along with the procedure followed in the administration and statistical evaluation of the Experimental Sentence Completion Test.

Chapter V will present the findings of the experiment and an analysis of those statistics. Inferences related to the hypotheses proposed will be drawn and discussed. Finally, a summary of the conclusions, and suggestions for further research will be attempted.

CHAPTER I

THE RATIONALE UNDERLYING SENTENCE COMPLETION TESTS

From a technique used initially as a means of assessing intellectual functions, the sentence completion method is now considered as one of the many so-called projective tools of personality assessment. This chapter presents a discussion of the rationale underlying the sentence completion technique. The concept of projection as applied to personality assessment is analyzed. This is followed by the logical application of the term to the sentence completion method.

1.- The Concept of Projection

Prior to 1939 projective methods were not designated as such; the term projective in connection with the methods known as such today, was introduced in an article by Frank¹. In that article Frank described the projective approach:

A projective method for the study of personality involves the presentation of a stimulus situation designed or chosen because it will mean to the subject not what the experimenter has arbitrarily decided it should mean (as in most psychological experiments using standardized stimuli in order to be 'objective') but rather whatever it must mean to

¹ L.K. Frank, "Projective Methods for the Study of Personality", in the Journal of Psychology, Vol. 8, 2nd Half, issue of October 1939, p. 389-413.

the personality who gives it or imposes upon it, his private, idiosyncratic meaning and organization.²

Frank apparently derived the term projective from another connected with the psychoanalytic literature -- projection. This term has numerous meanings,

some of which seem applicable as a partial description of the processes involved in these techniques, and some of which are unsuitable.³

Freud has described projection:

The projection of inner perceptions to the outside is a primitive mechanism which, for instance, also influence our sense-perceptions, so that it normally has the greater share in shaping our outer world. Under conditions that have not yet been sufficiently determined even inner perceptions of ideational and emotional processes are projected outwardly, like sense perceptions, and are used to shape the outer world, whereas, they ought to remain in the inner world.⁴

Warren defined projection as:

... the tendency or act of ascribing to the external world repressed mental processes which are not recognized as being of personal origin, as a result of which the content of these processes is experienced as an outer perception.⁵

2 Ibid.

3 John Elderskin Bell, Projective Techniques, a Dynamic Approach to the Study of the Personality, New York, Longmans, 1948, p. 1.

4 Sigmund Freud, "Totem and Taboo", in A.A. Brill, Edit., The Basic Writings of Sigmund Freud, New York, Random House, 1938, p. 857.

5 Howard C. Warren, Edit., Dictionary of Psychology, Boston, Houghton Mifflin, 1934, p. 212.

Whereas, Noyes⁶ referred to projection as a defence mechanism by which one disowned the unfavourable facets of one's own personality and saw them in other people, thus releasing feelings of tension (homeostatic mechanism).

Earlier, three well-known interpreters of psychoanalysis explained projection as

... a defensive process under sway of the pleasure principle whereby the Ego thrusts forth on the external world unconscious wishes and ideas, which, if allowed to penetrate into consciousness, would be painful to the Ego.⁷

Later Sears⁸ elaborated on this definition by explaining the psychoanalytic term projection as:

A wish, attitude, or habit-hierarchy which is not compatible with other attitudes or habits of an individual that may be attributed by that individual to other persons rather than to himself, providing he lacks insight into the fact that he himself possesses the trait in question. This process of attribution is unconscious, i.e., the subject does not give any verbal evidence that he knows his perception is false.

⁶ A.P. Noyes, Modern Clinical Psychiatry, Philadelphia, Saunders, 1934.

⁷ William Healy, Augusta Bronner and Anna Mae Bowers, The Structure and Meaning of Psychoanalysis, New York, Knopf, 1930, p.230.

⁸ R.R. Sears, "Survey of Objective Studies of Psychoanalytic Concepts," in the Social Science Council Monograph Bulletin, No. 51, 1943.

Bell⁹ pointed out that these psychoanalytic definitions of the process of projection have four points in common:

1. It is a defense mechanism against drives that are not conscious.
2. It is an unconscious process.
3. It results in unconscious impulses, feelings, ideas, and attitudes being attributed to others.
4. It serves to reduce an individual's tension.

Further, Sears made a distinction between those definitions as applied to the basic paranoid mechanism, and an implication of the term projective, as applied to the various techniques, wherein this term implies "that the motivational and organizational properties of a personality influence the perceptual and judgmental processes."¹⁰

The Latin roots of the word provide some clue to the action involved in those methods. Using the Latin then, projection can be considered as a "casting forward", in this case, the throwing outwards of the personality by some external behaviour. The technique being used merely draws out the personality picture, or as Frank¹¹ has suggested

⁹ Op. cit., p.2

¹⁰ R.R. Sears, op. cit.

¹¹ L.K. Frank, op. cit.

acts as a catalyzing agent to bring about the desired reaction.

Clinical psychologists, are, however, indebted to Rapaport for a more elaborate account of the projective process as applied to the various techniques encompassed by the term projective measures. He divorced the term projection from the psychoanalytic concept of the word, indeed, he separated it still further from

the popularized and emasculated version of that concept denoting any attribution of one's own intents, thoughts, and feelings to another person.¹²

Rapaport postulated the Projective Hypothesis, to which it is necessary to turn for a fuller explanation of the underlying psychological processes involved in projective testing.

"Projective procedures," Rappaport wrote, "imply a general 'projective hypothesis'."¹³ This was formulated as follows:

All behaviour manifestations of the human being, including the least and the most significant, are revealing and expressive of his personality, by which we mean that individual principle of which he is the carrier.¹⁴

12 David Rapaport, Merton Gill and Roy Schafer, Diagnostic Psychological Testing, the Theory, Statistical Evaluation, and Diagnostic Application of a Battery of Tests, Vol.2, Chicago, Year Book Publishers, 1946, p.6.

13 Ibid., p.6

14 D. Rapaport, "Principles underlying Projective Techniques," in Character and Personality, Vol.10, No.3, issue of March 1942, p.213-214.

Rapaport went on to include under the term behaviour,

... all of the following aspects: (a) behaviour historically viewed in the life-history; (b) behaviour statically viewed as reflected in the environment with which the subject surrounds himself, as the furniture of his house, the clothes he wears, etc.; (c) bodily manifestations, or patterns of voluntary, habitual and expressive movements; (d) internal behaviour, including percepts, fantasies, thoughts.¹⁵

In the projective process, the person plays an active and spontaneous role in the structuring of unstructured material. In doing that he reveals his structuring principles or the principles of his psychological make-up. Projective procedures draw forth such responses as play, perceptual - associative organizing, associative structuring, drawing or modelling which reveal the psychological essence of behaviour rather than an inventory of items about the personality.

Bell drew a parallel between the projective process in the test situation "and the projection involved in architecture or in map-making"¹⁶ namely, the reproduction of a three-dimension on the two-dimensional plane. Similarly, Rapaport¹⁷ suggested the parallel between the process and

¹⁵ Ibid., p.214

¹⁶ John Elderkin Bell, op. cit., p.3

¹⁷ David Rapaport, et al, op. cit.

the moving picture projection, where the projector represented the technique used; the film represented the personality; and, the picture seen was the test record. At the same time he emphasized that the more difficult the conditions of perceiving a given object, i.e., the more unstructured the situation, the more clearly the subject would stamp the percept with his own personality.

The Sufficient Criteria of Projective Tests. - In his earlier article, Rapaport suggested that there are many behaviour characteristics remote from the actual 'personality'; perhaps, because of a "cultural evolution resulting in the development of a hierarchy of functions, or to the conventionalizing effect of social patterns in our civilization."¹⁸ One of the sufficient criteria follows: "The realm of behaviour to which the projective technique is applied should not be a conventionalized one nor one remote from the core of the personality."¹⁹

By comparing personality pictures obtained in clinical practice with those obtained by the use of the various objectives four other sufficient criteria evolve:

18.D. Rapaport, op. cit., 1942, p.214

19 Ibid., p.214. Underlined in the original.

1. They must allow of easy objective observation.
2. They should allow of easy and objective registration.
3. The material obtained should be systemized or scored to permit intra- and inter-individual comparisons.
4. The subject should be unaware of the significance of the test and the reactions to it.

Later, Rapaport simplified the criteria²⁰ involved to:

1. The tests should be limited to one segment of behaviour.
2. They should be impersonal from the point of view of the tester.
3. They should be economical from the point of view of time.
4. They should be simple to give and easy to score.
5. They must, of necessity, be standardized, i.e., they should consist of the same materials and be given, where possible, by the same examiner, when given to different subjects.

These sufficient criteria obviously apply to any psychological test, hence, do not enable the psychologist to form any judgment as to whether or not a certain testing procedure constitutes a 'projective' or 'non-projective'

20 David Rapaport, et al., op. cit., p.8.

measure of personality. All that Rapaport has stated is that projective tests should meet certain criteria that should also be met in other psychological tests.

2.- Personality Theory and Projective Techniques

Most important of the theoretical constructs to be considered in the understanding of projective techniques is the conception of the term 'personality' as construed by the projective hypothesis. Since projective tests refer to the unconscious motivation of action and behaviour, this necessitates a personality theory that assumes the existence of, and accounts for, those motivations. Such a theory is found among those of the psychoanalysts. Two dangers are apparent here. First, it may be difficult to determine the extent of the conscious and unconscious, i.e., the various strata in the depth of the personality, and their influence in producing the test results. Second, psychoanalytical concepts may be applied to projective data haphazardly and regardless of independent thinking appropriate to the test material.

From the acceptance of a motivated theory of personality it follows that personality must be viewed as a dynamic not a static process. It is continually fluctuating in temporal space. This does not mean, however, that the personality is not a structured unit. The opposite must be

presumed, since the structure of the individual is brought about by the influence of the particular range of psychological, physiological, and sociocultural factors. Since behaviour is functional, "the personality structure, as well as the influence of the field in which the personality is operating, reveals itself in the behaviour of the individual" which "reflects the integral relationship between the demands of the self and the demands of the situation and is an attempt to adapt ^{to} these internal and external demands."²¹ It is important to note here the following remarks by the same author:

... the observable responses of the individual in specific situations are consistent with the personality in that situation, although they may not be consistent with other expressions of personality in other situations. Logical consistency of behaviour is not the same as psychological consistency, and, while the former may be present, the latter is always present. Since behaviour has this definite relationship to the personality structure, every act reveals the structure with some acts telling more than others about their origin. Projective techniques are attempts to order the most readily interpreted behaviour of the individual.²²

Rapaport summarizes the personality picture, as follows :

The subject has an Ego which is the recipient of outside stimulation, and which may be inclined to take, to shirk, or incessantly to invoke stimulations. This Ego is also the executor of the

²¹ John Elderkin Bell, op. cit., p.7.

²² Ibid., p.7-8.

intentions of the unconscious strivings, which in their particular constellation and strength are specific to the individual; as executor of these intents, the Ego may oppose them, subserve them without delay, or postpone them and by thinking prepare for their optimal realization. The Ego has a certain autonomy -- autonomous energy (bound cathexes), autonomous behaviour (defense mechanisms), and autonomous thought patterns -- to govern perception, execution and thought. The reception of stimulation by the Ego is not automatic but selective, and to some extent distorts the stimulation to meet the needs of the subject. The execution of the intentions by the Ego is likewise not an automatic discharge of internal tensions, but an adaption to the nature of the objects in reality which these intentions are aimed at or must cope with.²³

3.- Sentence Completion Tests as Projective Techniques

Various psychologists by introspection and by logical analysis have interpreted the sentence completion method in terms of the projective hypothesis. They have put forward certain theoretical assumptions upon which the technique is based. As will be seen later, little experimental work has been carried out to test these assumptions.

Hutt assumed partially structured personality tests, like the sentence completion method, to be projective techniques "because they elicit responses dependent in

²³ David Rapaport, et al., op. cit., p.11-12 underlined in the original.

part upon the projection by the subject of his personal interpretation or interaction into the stimulus".²⁴

Rohde suggested that questionnaires tended to place individuals on the defensive whereas projective techniques avoided such resistance. The latter, she considered, revealed latent needs, sentiments, attitudes and aspirations which the person undergoing the test would be unwilling or unable to recognize or express in direct communication. The sentence completion is an indirect approach in contrast with the direct approach used in questionnaire procedures. She described the sentence completion test as "essentially a projection technique utilizing free association".²⁵ In responding to the stimulus items the subject unconsciously reveals his true self, since there is no way in which he can anticipate the significance of his completions for personality study.

Stein considered that the clinician, in analyzing the responses to sentence completion items, accepts three theoretical assumptions:

- (1) when an individual is put under pressure to respond with the first idea that occurs to him,

²⁴ Max L. Hutt, "The Use of Projective Methods of Personality Measurement in Army Medical Installations", in the Journal of Clinical Psychology, Vol 1, No.2, issue of April 1945, p.136.

²⁵ Amanda R. Rohde, "Explorations in Personality by the Sentence Completion Method", in the Journal of Applied Psychology, Vol.30, No.2, issue of April 1946, p.169-170.

he usually offers significant material which he does not censor; (2) when faced with the problem of completing or structuring an unstructured situation an individual's response will be indicative of the true nature of his own reactions and sentiments; (3) in talking about others, an individual is apt to reveal himself.²⁶

In a text on projective psychology, Abt and Bellak wrote :

Sentence Completion Tests are based primarily on apperceptive distortion and analysis of the content of the completion words, inasmuch as the incomplete sentence technique often constitutes a social or emotional situation that the subject has to be prepared to interpret in terms of his established images and memories.²⁷

Trites, Holtzman, Templeton and Sells²⁸ considered the sentence completion method to be a projective technique of personality study in that it represents an ambiguous test situation designed to encourage the subject to invest the stimulus with his own wishes, impulses, fantasies and values. The subject will have projected certain aspects of his personality into the test response. Thus,

26 M.I. Stein, "Use of a Sentence Completion Test for the Diagnosis of Personality", in the Journal of Clinical Psychology, Vol.3, No.1, issue of January 1947, p.52.

27 Lawrence Edwin Abt and Leopold Bellak, Ed., Projective Psychology, Clinical Approaches to the Total Personality, New York, Knopf, 1952, p.73.

28.D.K. Trites, W.H. Holtzman, R. Templeton and S.B. Sells, Psychiatric Screening of Flying Personnel, Research on the SAM Sentence Completion Test, Project No. 21-0202-0007, Report No.3, Randolph Field (Texas), U.S.A.F. School of Aviation Medicine, 1953, p.1.

it has an advantage over questionnaire and other direct methods of personality study in that it appears to be less of a cross examination of one's character and so less provocative of deceit.

Rotter and Rafferty referred to the method as a semi-structured projective technique of personality study. They wrote:

As in other projective devices, it is assumed that the subject reflects his own wishes, desires, fears and attitudes in the sentence he makes.²⁹

The sentence completion test apparently, then, can be categorized with the projective methods classified by Saul Rosenzweig³⁰ as "apperceptive-dynamic". These procedures require the higher intellectual processes as are involved in the evolution of fantasies or artistic productions. Interpretation reveals themes, drives, urges, or motives that are basic to the behaviour of the subject.

This chapter has presented a discussion of the theoretical assumptions underlying the use of sentence completion techniques of personality assessment. In the chapter which follows a review of the use of the technique to

29 Julian Rotter and Janet E. Rafferty. "The Rotter Incomplete Sentence Blank", in Arthur Weider, Ed., Contribution toward Medical Psychology, Theory and Psychodiagnostic Methods, Vol.2, New York, Ronald Press, 1953, p.591.

30 Saul Rosenzweig, "Investigating and Appraising Personality", in T.G. Andrews, Ed., Methods of Psychology, New York, Wiley, 1948, p.560-564.

1957 will be undertaken, leaving those studies which have a direct bearing on the types of completion items used to a later chapter.

CHAPTER II

GENERAL STUDIES OF THE SENTENCE COMPLETION METHOD

Considered as a method, which evolved out of the word association lists, the sentence completion technique has a history dating almost to the beginnings of modern experimental psychology itself. Like most other so-called projective techniques its use preceded its validation and certainly its rationale. This chapter presents a review of the use of the technique; its use up to 1947 - the year which saw the publication of the first reports on the Rotter Incomplete Sentence Blank, one of the now more widely used versions of the method; and, those contemporary studies from 1948 to 1957. Studies relating to the use of different types of items are reviewed in the subsequent chapter.

1.- Literature to 1947

The use of incomplete sentences as items for a psychological test dates back to 1897 when Ebbinghaus used sentence completions for the purpose of studying mental capacity and reasoning ability.

The first report of the use of this type of test for eliciting inhibited personality traits was that of

Payne¹ at the New York Guidance Clinic in 1928. He used a 50-item test as one of a battery of guidance tests in the College of the City of New York, the Hebrew Orphan Asylum in New York City, in other institutions, and in industrial settings.²

In 1930, Tendler³ published a preliminary -- and apparently final -- report on his Test for Emotional Insight. This test consisted of twenty incomplete sentences. All were self-referent items constructed by Tendler to determine "trends, fixed attitudes, attachments to persons, conflicting desires, satisfactions, and annoyances",⁴ The criteria used in selecting such items were that they should directly evoke an emotional response, allow the subject freedom of response, and not provoke discrimination or choice by the subject.

The stimuli include a major share of important emotional states; admiration, anger, happiness, love, hate, self abasement, worried state,

1 Arthur F. Payne, Sentence Completions, New York, Guidance Clinic, 1928.

2 Amanda R. Rohde, The Sentence Completion Method, Its Diagnostic and Clinical Application to Mental Disorders, New York, Ronald Press, 1957, p.7.

3 A.D. Tendler, "A Preliminary Report on a Test for Emotional Insight", in the Journal of Applied Psychology, Vol.14, No.2, issue of April 1930, p.122-136.

4 Ibid., p.122

compensatory make-believe, regret, boastfulness, pride, grudge formation, negativism, pity, shame, fear, interest, disgust, evasion, desire.⁵

In his report, Tandler presented a tabulation of ten subjects' responses to the test. These ten were obtained from a sample of 250 college girls, seventy-five per cent of whom were between eighteen and twenty years of age. These were also given the Woodworth Personal Data Blank. They were listed in order of increasing Personal Data unfavourable scores -- from 0 to 41 out of seventy items. The Emotional Insight Test blank for every twenty-fifth individual was drawn out. These constituted the ten examples given.

In his quantitative validation study, Tandler administered the Otis SA Higher Examination, the Woodworth Personal Data Blank, and his Test for Emotional Insight to 190 young women aged from seventeen to twenty-seven years, seventy-six per cent of whom were between eighteen and twenty years.

Using the Personal Data Blank scores, the group was divided into four quartiles, the first and last being chosen as representing favourable and unfavourable emotionally organized groups. Comparison of these two groups on the Otis

⁵ Ibid., p.124.

and on the Test for Emotional Insight was made. The two groups were not significantly different in intelligence but there was a significant difference between the groups on the Test for Emotional Insight and on the Woodworth Personal Data Blank, the former being somewhat more discriminative. In his discussion of the results, Tendler noted that a stimulus item evoked different responses from different persons.

Tendler's subjects also submitted autobiographical sketches. These were compared qualitatively with the sentence completions. Tendler concluded that clinically his test was found to be of value in eliciting attitudes, trends, and significant clues. The results of his study were not cross-validated.

Cameron⁶ studying the nature of the language and the thinking of senile and of schizophrenic patients in contrast with normal individuals, utilized a 15-item completion test. He gave the test to four groups: twenty-nine normal children, twenty normal adults, twenty-two deteriorated seniles and twenty-five disorganized schizophrenics. The test was given orally to each individual who was expected to give an oral completion.

⁶ Norman Cameron, "A Study of Thinking in Senile Deterioration and Schizophrenic Disorganization", in the American Journal of Psychology, Vol.51, No.4, issue of October 1938, p.650-664; also, "Reasoning, Regression and Communication in Schizophrenia", in Psychological Monographs, Vol.50, No.1, issue of 1938, p.1-34.

From a qualitative analysis, Cameron concluded that the groups could be differentiated on the basis of their responses. His report went into the thought processes of the abnormal groups. One outstanding conclusion was that schizophrenic thought processes do not resemble those of children as many authors had suggested.

A decade later Cameron and Margaret⁷ reported on an orally-given sentence completion test in an investigation of the correlates of scattered speech in the responses of normal adult subjects.

A second quantitative study of the validation of sentence completion items, reported in 1941, was that of Lorge and Thorndike.⁸ The 240 items used were of several types but tended to be relatively impersonal. Their test was administered to a group of normal subjects who were asked to respond to each item as rapidly as possible with just a single word. Each author attempted to assign, independently of the other, each response to one of seventy

7 Norman Cameron and Ann Margaret, "Experimental Studies in Thinking I, Scattered Speech in Response of Normal Subjects to Incomplete Sentences", in the Journal of Experimental Psychology, Vol.39, No.5, issue of October 1949, p.617-627.

8 I. Lorge and E.L. Thorndike, "The Value of the Responses in a Completion Test as Indications of Personal Traits", in the Journal of Applied Psychology, Vol.25, No.2, issue of April 1941, p.191-199.

personality variables. Each response was then given a weight of plus 1 or minus 1 for strength or weakness respectively in the variable to which it was assigned. A total score for each variable was then found by adding algebraically the weights assigned to the items scored for that variable.

Using data obtained by self-ratings and by direct questioning, the authors investigated the validity of the two most reliable categories, the majority of the scoring reliabilities being very low. The correlations obtained were around zero. From this Lorge and Thorndike concluded that the verbal replies in association and sentence completion tests were unrealistic in terms of attitudes elicited.

Several criticisms of that study have been voiced. In the first place, the test was given as a speed test. Secondly, the subjects were restricted to one-word responses. Thirdly, as has been noted above, the reliabilities of the scoring procedure computed by the split-half method, were very low. Rohde⁹ criticised the wording of the items themselves as well as their selection, whilst Bell¹⁰ pointed out that the results may have depended upon the type of stimulus

⁹ Amanda R. Rohde, op. cit., p.10

¹⁰ John Elderkin Bell, op. cit., p.46-47.

items used and upon the techniques of interpretation.

Sanford¹¹ described a 30-item Completion Sentences Test that had been assembled for use in clinical situations, with a view to obtaining measures for a need-press analysis based on Murray's Personology. One type of item used stated a press object in order to determine an individual's reaction to such presses. Another type stated an action pattern of some need object to elicit what press the subject would associate with particular need-action patterns. A third, consisted of ambiguous statements to which the subject would supply needs.

Sanford gave the test to school children in grades three to nine, along with a Completion of Pictures Test and an Interpretations Test. The numbers in each grade were small. He presented no quantitative analysis of the study, although it was concluded that the qualitative results showed differences in personality among those tested.

In 1941, Rohde and Hildreth published a revision of Payne's test. This revision contained sixty-four items in place of the fifty used by Payne. After research carried over several years, Rohde presented the results in an

¹¹ R.N. Sanford and E.A. Cobb, "Studies of Personality and Environment, Physique, Personality, and Scholarship", in Monographs of Social Research in Child Development, Vol. 8, No.1, issue of 1943, p.1-705.

article published in 1946,¹² and more recently in a book on the Rohde Sentence Completion Method.¹³

The validation study of the Rohde-Hildreth Test was carried out with fifty boys and fifty girls selected at random from 670 ninth-grade high school students. The subjects were required only to complete the sentences. Those were then scored in terms of the need, press or inner states indicated. Over-all ratings were made for each of thirty-three variables on a ten-point scale. The ratings were based on the judgment of the experimenter. To obtain criterion data, the subjects were interviewed by teachers and counsellors who then rated each boy or girl interviewed. Corrected Pearson product-moment correlations for all the variables combined were .79 for the girls and .82 for the boys. Most of the coefficients for individual variables were between .80 and .90. Reliability coefficients were found by retesting twenty-one girls and twenty-three boys. These coefficients were found to be, respectively, .82 and .76.

12 Amanda R. Rohde, "Explorations in Personality by the Sentence Completion Method", in the Journal of Applied Psychology, Vol.30, No.2, issue of April 1946, p.169-181.

13 Amanda R. Rohde, op. cit., xii-30lp.

Seventy-eight per cent agreement was found on the scoring of individual items for the same twelve papers by four judges. Since the rater's over-all judgment of each test was dependent also upon frequency and intensity, it is presumed that this index of inter-scorer reliability does not cover those over-all ratings.

Shor, in 1946, reported on the Self-Idea Completion Test.¹⁴ The fifty incomplete sentences described were mainly of the self-referent type, in that many items contained "I" or "My". Subjects were encouraged to be free and spontaneous in their responses. The items themselves were suggestive of feeling tones, qualities of attitude and specific objects or areas of attention. In interpreting responses the examiner used his clinical judgment utilizing such gross variables as areas of rejection, areas of resistance, and methods of evasion. Whilst Shor's test was intended for use in military psychiatric installations, he stressed that the test should be designed for a specific situation. He also suggested that attention should be paid to the sequence and the distribution of items, the planting of neutral or shock absorbing items, and the arrangement of items. No quantitative data were given.

¹⁴ J. Shor, "Report on a Verbal Projective Technique", in the Journal of Clinical Psychology, Vol.2, No.1, issue of February 1946, p.279-282.

Using a modified version of Tendler's Test for Emotional Insight, Carter¹⁵ investigated the relationship between the various stimulus items presented visually, one by one, and the galvanic skin responses (GSR). Carter labeled the orally-given sentence completions into ego-positive, ego-negative, social positive and social negative, fear and aversion, and unclassified. In his paper, he reported that, in general, the GSR, as well as the reaction time, discriminated a group of twenty persons who reported no maladjustments, from a group of twenty persons who reported some maladjustments and twenty individuals receiving psychiatric aid. The palmar skin resistance and prolonged reaction time also indicated the area as well as the intensity of emotional disturbances.

On the sentence completions, however, the three groups varied little except for responses classed as "social positive" and "unclassified". The neurotic had fewer "social positive" and more "unclassified" sentence completions than did the normals.

¹⁵ Homer L.J. Carter, "A Combined Projective and Psychogalvanic Response Technique for Investigating Certain Affective Processes", in the Journal of Consulting Psychology, Vol. 11, No. 5, issue of September-October 1947, p. 270-275.

The Rotter Incomplete Sentences.- Rotter and Willerman¹⁶ adapted items from tests used by Hutt, Shor and Holzberg on a clinical basis with the intent to avoid using items that elicited stereotyped responses. At the same time, the items selected were as little structured as possible. Their 40-item test contained both self-referent and impersonal items. The test was developed for use in Army Air Force Convalescent Hospitals and later was adapted for the screening of maladjusted college students. The scoring technique adopted was as objective as possible, so that persons scoring required no complex training. Responses were compared with a reference list of examples, and rated along a seven-point scale. Thus, each response could have been given a numerical score of from plus 3 to plus 1 for conflict or unhealthy responses, zero for neutral responses, and from minus 1 to minus 3 for positive or healthy responses. The example used for comparison in the scoring manual were selected on the basis of two criteria: small samples from various types of psychiatric and non-psychiatric Air Force Hospital patients; and, the experience of the

¹⁶ J.B. Rotter and B. Willerman, "The Incomplete Sentence Test as a Method of Studying Personality", in the Journal of Consulting Psychology, Vol.11, No.1, issue of January-February 1947, p.43-48; and, "Development of a New Projective Test, Incomplete Sentence Test DE3034", in S.W. Bijou, Ed., The Psychological Program in Army Air Force Convalescent Hospitals, Army Air Forces Aviation Psychology Research Report No.15, Washington, Superintendent of Documents, 1947.

authors and their observations on the general principles of mental health. The samples of patients used to establish scoring standards were not included in the subsequent studies.

Rotter and Willerman divided 200 patients in U.S. Air Force Convalescent Hospitals into three groups using all available information except the sentence completion test results: the psychologically fit who could return to duty at once, the psychologically unfit for immediate duty but who should benefit from hospital treatment, and those too disturbed for further Air Force service or who were recommended for psychiatric hospital intake.

The authors reported a reliability index of .85 using the split-half method corrected by the Spearman-Brown prophesy formula. Seven judges rated the same fifty records and the average interscorer reliability coefficient was found to be .89. Using as a criterion an estimate by clinical psychologists of the severity of the psychological disturbance, the validity of the test expressed by a triserial correlation was found to be .61.

The test was administered to 148 patients at another convalescent hospital. The tests were rated by experienced clinical psychologists on a scale of "no", "mild", "moderate", and "severe" disturbance. Inter-scorer reliability was .68, lower than that found by using more objective methods. Biserial correlations of .41 for one

rater and .39 for a second rater were found by using as a criterion an admitting diagnosis of 'psychiatric' or 'non-psychiatric'.

The authors concluded that the sentence completion test used was effective for the purpose of screening patients into rough categories as to type and severity of disturbance.

Rotter, Rafferty and Schachtitz¹⁷ described their use of the Rotter Incomplete Sentences, an adaption of the Army Test, as a device for the screening of college students. The scoring method used was comparable to that used with the Army Air Force Convalescent Hospital patients, but separate scoring manuals were drawn up for male and female groups. The reliability and validity indices were computed for both sexes. Interscorer reliability coefficients based on fifty males and fifty females were .91 and .96 respectively. The reliability indices based on 124 men and seventy-one women college students were .84 and .83 respectively.

Using as the criterion, ratings of "adjusted" or "maladjusted" given by the college instructors, validity

17 Julian B. Rotter, et al., "Validation of the Rotter Incomplete Sentence Blank for College Screening", in the Journal of Consulting Psychology, Vol.13, No.5, issue of October 1949, p.348-356. Also Julian B. Rotter and Janet E. Rafferty, Manual, the Rotter Incomplete Sentences Blank, College Form, New York, Psychological Corporation, 1950, 86p.

indices of .62 and .50 were obtained on seventy-eight male students and seventy-two female students respectively.

Morton¹⁸ used the Rotter Incomplete Sentences Blank as one of several measures of the effectiveness of brief psychotherapy. The Incomplete Sentences Blank correlated .53 with judges' ratings of current psychological status and .40 with the Mooney Problem Check List. Using the Incomplete Sentences, Barry¹⁹ obtained a correlation of .67 between it and the ratings of terminal counselling interviews.

A special manual was developed by the staff of the Human Resources Research Laboratories²⁰ for scoring completions made to the Incomplete Sentences Blank. The purpose of the manual was to yield one general score which could be used as a predictor of an individual's adjustment

18 Robert Morton, A Controlled Experiment in Psychotherapy Based on Rotter's Social Learning Theory of Personality, unpublished Doctoral thesis presented to the Ohio State University, Columbus, 1949.

19 J.R. Barry, "The Relation of Verbal Reactions of Adjustment Level", in the Journal of Abnormal and Social Psychology, Vol.45, No.4, Issue of October 1950, p.647-658.

20 Human Resources Research Laboratories, Studies in Air Crew Composition, Study No.2, The Use of the Rotter Incomplete Sentences Blank in the Prediction of Air Crew Adjustment, Part 2, Washington D.C., Headquarters Command U.S.A.F., 1952, ix - 18p.

to membership on a combat air crew in the U.S. Air Force. Scoring principles were developed so as to classify the specific completion into a general category and, at the same time, to discriminate between "high" and "low" criterion groups. The manual, in its final form, contained the scoring principles developed for each stem, examples of specific items scored under each principle, and general scoring instructions.

Using a trial manual for the first six stems the inter-scorer reliability for two scorers was found to be .79. The same two scorers using Rotter's Manual to score the first six stems of the same sample of fifty subjects correlated .70. Using the final form of the manual with twenty-two subjects, inter-scorer correlations for four scorers ranged from .69 to .91.

The reliability of the test based on a sample of seventy subjects, using odd versus even stem halves and corrected by the Spearman-Brown formula, was found to be .60. In an attempt to obtain equivalent halves by matching stems by pairs a corrected reliability coefficient of .46 was obtained.

A validity study using the new manual and a different sample of fifty-one aircrew officers and twenty-three airmen of aircrew status produced a validity coefficient of .40. The criterion score of aircrew adjustment was computed for each individual using data from self and peer ratings.

Five studies were reported on the predictive validity of the Incomplete Sentences. The authors concluded:

The completions which crew members made to the stems of this blank, at the time of crew training, were related to their subsequent performance in combat.²¹

Rotter, Rafferty and Lotsof²² later reported on a study concerned with the development of a scoring manual, and with a preliminary test of the validity, for the High School Form of the Incomplete Sentences Blank for assessing pupils as adjusted or maladjusted.

From case material collected, the authors selected examples of completions for each item to represent various points on a 7-point scale. Separate manuals were constructed for boys and girls. These manuals were then tried out on a new sample of tests to determine their adequacy and comprehensiveness. Cross-validation was then carried out on further groups of boys and girls.

Using an odd-even estimate of reliability corrected by the Spearman-Brown formula reliability coefficients of .74 and .86 were found with groups of fifty boys and

²¹ Ibid., p.18. Underlined in the original.

²² J.B. Rotter, J.E. Rafferty and A.B. Lotsof, "The Validity of the Rotter Incomplete Sentences Blank, High School Form", in the Journal of Consulting Psychology, Vol.18, No.2, issue of April 1954, p.105-111.

fifty girls respectively. Interscorer reliability for trained scorers was reported as .96 on the boys' manual and .97 on the girls' manual.

In the validity studies, the criterion of degree of adjustment-maladjustment was based on interviews given about a year from the giving of the test, as well as on sociometric ratings. With a group of forty-five boys the Pearson r between interview ratings and the Incomplete Sentences' scores was .20 with a Standard Error of .14. With a group of forty-eight girls, the Pearson r between interview ratings and the Incomplete Sentences' scores was .37 with a Standard Error of .12. Using a group of sixty-eight boys, the sociometric scores correlated .20 with the Incomplete Sentences' scores, with a Standard Error of .12. Using seventy girls, the coefficient of correlation between the same two measures was .32, with a Standard Error of .11.

In the first half of this chapter the studies of sentence completion tests up to 1947 have been surveyed. In some cases where later studies were reported dealing with tests developed mainly in this period, these have been included to provide continuity. The subsequent half of this chapter will scan the years 1948 to 1957.

2.- Literature from 1948 to 1957

From 1948, the sentence completion technique appears to have been generally accepted as a reliable and valid tool for investigating personality, so that the literature from that year on was devoted to the use of the device rather than to its validity and reliability.

Kline²³ reported on a study of sex differences on a 10-item sentence completion test. Each of the stimulus phrases was presented orally and the subject had fifteen seconds in which to respond before his response was timed as being over the normal limits. Side remarks made by the subject were also noted. In the study, Kline used two hundred males and one hundred females between the ages of seventeen and thirty-eight years of age. Each completion test was rated on (1) the reaction time; (2) long pauses or "no response" items; (3) the length of the completions; (4) behavioural manifestations indicative of tension, anxiety, self-consciousness, etc.; (5) contaminations due to corrections or alterations; (6) intensity of language used. If four or more items were significant, the test was classed as positive. From the study, Kline concluded that sex differentiation is distinct. Costin and Eiserer have

23 Milton Kline, "A Short Form Sentence Projection Technique", in the Journal of General Psychology, Vol.39, Second Half, issue of October 1948, p.273-287.

described a study to assess students' attitudes towards various phases of school life. They utilized an 8-item test containing impersonal items structured around school situations. The sample studied consisted of seventy-four eleventh grade pupils, equally divided into males and females. The 786 completions were each classified according to attitudes expressed towards school in general, specific aspects of school life, teachers and students. Each response was rated further as "positive", "negative" or "neutral". The following conclusions were stated:

1 Students tended to identify with one another more than they did with other aspects of school life.

2 Students expressed a "stereotyped opposition attitude" towards school as an institution. This "opposition" was not apparent toward specifics of school life.

3 There were more negative attitudes expressed with regard to teachers than with reference to any other area.

4 The sentence completion test was a suitable instrument for investigating "school morale".²⁴

Hadley and Kennedy²⁵ at Purdue University, compared the performance on the first part of the OSS Sentence

24 F. Costin and P.E. Eiserer, "Students' Attitudes toward School Life as Revealed by a Sentence Completion Test", an abstract in the American Psychologist, Vol.4, No.7, issue of July 1949, p.289.

25 John M. Hadley and Vera E. Kennedy, "A Comparison between Performance on a Sentence Completion Test and Academic Success", in Educational and Psychological Measurement, Vol.9, No.4, Winter Issue of 1949, p.649-670.

Completion Test and the academic success of 157 college students enrolled in an Introductory Psychology course. As a criterion measure they used the relationship between the percentile score on the American Council on Education Psychological Examination for College Freshmen, and the percentile level of grade-point average. Two groups were used. Individuals in each had A C E percentile scores of sixty or greater, but those in the first group had grade-point percentiles of forty or lower, whereas, those in the second group had grade-point percentile scores of sixty or greater. After scoring each completion as conflictual, positive or neutral, a total score for each individual was calculated by subtracting the total negative responses from the total positive responses.

The authors found that the subjects who achieved in accord with their ability had less tendency to experience conflict than the underachievers. A Critical Ratio for the difference of the means of the two groups was significant at the .04 level.

Hadley and Kennedy noted a relatively high inter-scorer reliability coefficient and a high rescore reliability coefficient when the same scorer repeated the task. The reliability of the test was estimated by the odd-even method.

Karen Machover²⁶ developed a sentence completion test for the behavioural study of the reactions of children at the Caroline Zachry Institute, Kings County Hospital, New York. Case reports were issued the same year.²⁷

Wilson²⁸ made use of a sentence completion test to discriminate between well-adjusted and mal-adjusted secondary school pupils. She used twenty-two subjects in all. The maladjusted group had more boys than girls, and the well-adjusted group had more girls than boys. The author concluded that the formal aspects of the completions showed no differences between groups. It was noted, however, that there were more omissions among the well-adjusted than among the maladjusted. Some discriminations were also reported on specific items.

This study was typical of many of the studies in this area, namely, in that it was not well designed. The

26 Karen Machover, CZI Sentence Completion Test, Brooklyn, New York, King's County Hospital, 1949.

27 Solomon Machover, Ed., Case Reports in Clinical Psychology, Brooklyn, New York, King's County Hospital, Vol.1, No.1, issue of August 1949.

28 Isabel Wilson, "The Use of a Sentence Completion Test in Differentiating between Well-Adjusted and Mal-Adjusted Secondary School Pupils", in the Journal of Consulting Psychology, Vol.13, No.6, issue of December 1949, p.400-402.

criteria of adjustment-maladjustment were not specified. The groups studied were small and the mixing of the sexes was unfortunate.

In a comparison of the attitudes expressed toward the father and the mother by psychoneurotic patients in the Stein Sentence Completion Test with those expressed elsewhere, Rosenberg²⁹ used a group of seventy-two male, white, veterans. The sentence completions dealing with attitudes toward parental figures were scored positive, negative or evasive. These were then compared to data from the patients' case history files. The attitudes expressed toward the father were in agreement in fifty-eight per cent of the cases; those expressed toward the mother were in agreement in sixty-nine per cent of the cases. Evasive responses were given in thirty-one per cent of the cases -- for attitudes toward the father -- and in seventeen per cent for attitudes toward the mother. The author concluded that the test is potentially a valuable tool in giving clues to the attitudes of a patient towards his parents.

In an extensive study of the predictive validity of a battery of projective tests that included a sentence

29 Selig Rosenberg, "Some Relationships between Attitudes Expressed toward the Parent in a Sentence Completion Test and Case History Data", in the Journal of Projective Techniques, Vol.14, No.2, issue of June 1950, p.188-193.

completion test, the Thematic Apperception Test, the Rorschach and the Bender Visual Motor Gestalt, for the selection of clinical psychologists in the Veterans' Administration Training Program, Kelly and Fiske³⁰ reported low positive validity coefficients for all four tests. The sentence completion test used, however, had the highest coefficient.

Fifty other-referent items were used by Kimball³¹ in her study of the relationship between personality factors and scholastic achievement. The subjects of the study were twenty adolescent boys of high intelligence but who were failing in school work. As a control group, Kimball used one hundred students who were not failing in school. Two hypotheses were expounded:

1. A significantly higher number of the under-achievers would reveal an essentially negative relationship with the father than would be found in the total population.

2. Aggressive feelings would be a source of guilt and anxiety more frequently among the underachievers than in

30 E.L. Kelly and D.W. Fiske, The Selection of Clinical Psychologists, Ann Arbor, University of Michigan, 1948; "The Prediction of Success in VA Training Program in Clinical Psychology", in the American Psychologist, Vol.5 No.8, issue of 1950, p.395-406; and, The Prediction of Performance in Clinical Psychology, Ann Arbor, University of Michigan, 1951.

31 Barbara Kimball, "The Sentence Completion Technique in a Study of Scholastic Underachievement", in the Journal of Consulting Psychology, Vol.16, No.5, issue of October 1952, p.353-358.

the total population and the former would be less able to give direct, effective expression to these negative feelings than would the members of the randomly selected group from the total population.

Of the total of fifty items only seven were apparently classified as positive, negative, or neutral. Two of these referred to the father; two had reference to guilt and anxiety feelings; the remainder had reference to the way in which aggression would be expressed. The author concluded that the experiment supported her hypotheses.

The Curtis Completion Form was stated to have been "designed to evaluate emotional maturity and adjustment".³² It consists of fifty incomplete sentences permitting free association, and two structured items. Forty items are of the self-referent type. Scoring is objective in that completions are scored 2 points, 1 point, or zero, depending on their content. The Manual³³ contains examples of each type of score. The author stated that the final form of the test evolved from eighty-four items. The standardization data were obtained on 175 normals, sixty neurotics and one

³² James W. Curtis, Examiner Manual for the Curtis Completion Form, Form A, Chicago, Science Research Associates, 1953, p.1.

³³ Ibid., p.2-4

hundred psychotics. The reliability of the test, estimated by the use of the split-half technique, was based on a sample of 133 adults. The estimated coefficient was .829 with a Standard Error of .027. Inter-scorer reliabilities ranged from .89 to .95, the highest coefficient being obtained when the two scorers were both psychologists. No cross-validation data were reported.

Luft, Wisham and Moody³⁴ used thirty-four sentence beginnings to elicit attitudes of neuropsychiatric, gastrointestinal, cardiovascular and surgical patient groups to the hospital staff, hospital conditions, other patients and their own ailments. The completions were scored positive, negative or zero with respect to the attitudes expressed. It was reported that different groups tended to have different attitudes.

The responses of eighty-three United States Naval Cadets to a series of sentence beginnings were utilized by Izard, et al.,³⁵ as a basis for constructing a multiple-choice sentence completion test. The latter consisted of

34 Joseph Luft, Wayne Wisham and Hamilton Moody, "A Projective Technique to Measure Adjustment to Hospital Environment", in the Journal of General Psychology, Vol.49, Second Half of October 1953, p.209-219.

35 Carroll Izard, et al., Construction and Validation of a Multiple-Choice Sentence Completion Test, an Interim Report, Project Number NMO01 077.01.02, Report No.2, Pensacola, U.S. Naval School of Aviation Medicine, 1953, 9 - xp.

forty-seven stems, each with a choice of three completions. A scoring key for thirty-two of the items was developed by having three psychologists rank the multiple-choice completions with respect to the degree to which those responses indicated inadequacy or maladjustment. Using as criteria the self-, and peer-ratings on the variable "social adequacy" for a group of 335 Naval Cadets validity coefficients of minus .15 and minus .22 were obtained respectively. These correlations were in the predicted directions and were significant at the .01 level of confidence. On test-retest with a group of 137 subjects, a product-moment reliability coefficient of .73 was obtained. In view of the data collected, the authors expressed the opinion that further development of the instrument was indicated, particularly since scoring of the completions was completely objective and could, therefore, be done by personnel untrained in psychodynamics.

Two exponents of the clinical approach to personality appraisal, as distinct from the psychometric approach, have suggested the value of sentence completions is that they provide "material which substantiates or clarifies very subtle indications from other techniques",³⁶ and add greatly

³⁶ James Quinter Holsopple and Florence R. Miale, Sentence Completion, a Projective Method for the Study of Personality, Springfield, Charles C. Thomas, 1954, p.5.

to a projective battery where TAT and Rorschach protocols are meager. Their book, which is essentially a manual for the Miale - Holsopple Sentence Completion³⁷ technique, contains fourteen illustrative records. The authors' objective "was to obtain material from which we might draw valid inferences concerning unconscious and semi-conscious desires, motives, conflicts, and systems of personality organization".³⁸

In a study of additional changes among college freshmen over a four-month period, Lingren³⁹ used a 20-item sentence completion device. His study was confined to a group enrolled in a first semester course in psychology. Completions were classified as "acceptance", "rejection" or "neutral" with regard to elements presented by the stimulus items. Differences in attitude and change of attitude were noted when results were compared according to the sex of respondents and when the results of the first administration were contrasted with the second administration. The author expressed the feeling that the test, while useful as a research tool, was too tedious and time consuming to score for use as an evaluational tool in routine use.

37 Ibid., xiii - 177 p.

38 Ibid., p.11

39 Henry Clay Lindgren, "The Use of a Sentence Completion Test in Measuring Attitudinal Changes among College Freshmen", in the Journal of Social Psychology, Vol.40, First Half of August 1954, p.79-92.

A novel use of sentence completions was reported by Gekoski and Izard,⁴⁰ who used the sentence completion method for item collection for questionnaires. The authors put forward three advantages: 1. The yield of items is great; 2. The items have the desired orientation; and, 3. The items are worded in the manner of the population with which they are to be used.

A unique study of sentence completions has been reported by Trites,⁴¹ who used an 88-item test to evaluate two assumptions often made in the objective interpretation of that type of test. Those assumptions may be stated:

1. Incomplete sentence stimulus items whose verbal content and meaning are immediately apparent and generally agreed upon will elicit responses which refer to that generally agreed upon meaning.

2. Systems of response classification which represent the attitudes and objects referred to by both stimulus and response may be developed by considering only the verbal content of the stimuli. In other words, it is assumed that

⁴⁰ Norman Gekoski and Eleanore S. Izard, "Note on Another Use of the Sentence-Completion Technique", in the Journal of Applied Psychology, Vol.3, No.2, issue of April 1955, p.139.

⁴¹ David K. Trites, Psychiatric Screening of Flying Personnel, Evaluation of Assumptions underlying Interpretation of Sentence Completion Tests, Report No.55-33, Randolph Field (Texas), U.S.A.F. School of Aviation Medicine, 1955, 9 p.

response to the items refer to their most common meaning, and that an adequate scoring system based on categories of item meaning can be constructed.

A factor analysis of the inter-correlations of seventy-four of the eighty-eight items yielded partial support for both assumptions. To obtain the data for the computations of inter-item correlations, each response of 392 aviation cadets and student officers was classified dichotomously as indicating either a positive or negative attitude, with reference to adjustment to the pilot training program. From that dichotomous scoring, inter-item tetrachoric correlations were computed. Thus, 2,701 inter-correlations constituted the basic matrix for factoring by the Thurstone complete centroid method.

Four factors were produced: the Self-Centred Anxiety Factor, the Air Force Motivation Factor, the Interpersonal Factor, and the Narcissistic (or Self-Enhancement) Factor. Trites concluded:

The four factors do indicate that it is possible to construct tests of this type to elicit responses relative to specified areas, and that these areas do have some reality in terms of stimulus - response content which transcends the arbitrary classification of the test constructor.⁴²

⁴² Ibid., p.9.

Burwen, Campbell and Kidd⁴³ reported a correlation coefficient of .32 between a sentence completion test designed to measure attitudes towards superiors and subordinates, and a direct attitude measure of the same dimension. A correlation coefficient of .27 was obtained between the completion test and an indirect measure based on an information test. For the experiment the authors used a group of 312 U.S. Air Force cadets in advanced training. Both correlations are significant at the one per cent level of confidence.

A report, which goes back to the initial use of the sentence completion technique, i.e., measuring intelligence, was issued by Copple⁴⁴. He utilized sixty items only thirty of which were given to any one subject. Items numbered one to thirty were for use with the age group five to eight years (Level A); items sixteen to forty-five were for use with ages nine to twelve years (Level B); and, items thirty-one to sixty were given to ages thirteen years and over -- including adults (Level C). The author reported a median interscorer reliability of .97 for ten scorers,

⁴³ Leroy S. Burwen, Donald T. Campbell and Jerry Kidd, "The Use of a Sentence Completion Test in Measuring Attitudes towards Superiors and Subordinates", in the Journal of Applied Psychology, Vol.40, No.4, issue of August 1956, p.248-250.

⁴⁴ George E. Copple, "Effective Intelligence as Measured by an Unstructured Sentence Completion Technique", in the Journal of Consulting Psychology, Vol.20, No.5, issue of October 1956, p.357-360.

and split-half reliabilities for three studies of .94, .95 and .93. Three methods of validating the test were tried, and the results were summarized thus:

1. Copple Sentence Completion Test scores compared favourably with the Binet and Goodenough mental ages in their correlation with chronological age.

2. CSCT scores compared favourably with Binet mental ages in predicting Stanford Achievement Test scores.

3. CSCT scores were significantly related to IQ gains subsequently made by a group of retarded children undergoing play therapy.

Fifty offenders and fifty non-offenders were matched in age, education, and length of naval service. Locke⁴⁵ reported that both groups were given the Stein Sentence Completion Test. The test forms were coded and mixed together. The hundred forms were then rated on eleven areas by three experienced clinical psychologists. Four categories for rating responses were used: "severely disturbed", "mildly disturbed", "no significant disturbance", and "insufficient evidence". A general clinical impression of each protocol was also expressed.

⁴⁵ Bernard Locke, "Comparison of Naval Offenders with Non-Offenders on a Projective Sentence Completion Test", in the U.S. Armed Forces Medical Journal, Vol.8, No.12, issue of December 1957, p.1825-1828.

The author reported differences between the prisoners and the control subjects in the expected direction in attitude toward inferiors, attitude toward peers, guilt feelings and attitude toward own abilities. These differences were significant at the .01 level of confidence. The groups were different in energy level and by general clinical impression at the .05 level of confidence; the differences being in the expected direction. Thus, it was possible for the psychologists to separate prisoners from non-prisoners on the basis of the test protocols with a better than chance success. However, the control group scored higher than expected on at least four of the areas under consideration; attitude toward family unit, attitude toward superiors, fears, guilt feelings and on the general clinical impressions. The author hypothesized that men volunteer for military duty because of maladjustment in those areas, as an explanation for those findings.

Masling⁴⁶ arranged appointments for eight members of a class in projective techniques to administer and interpret two Rotter Incomplete Sentences tests. The examiners were not aware that the test subjects had memorized completions which were so constructed as to be considered

⁴⁶ Joseph M. Masling, "The Effects of Warm and Cold Interaction on the Interpretation of a Projective Protocol", in the Journal of Projective Techniques, Vol.21, No.4, issue of December 1957, p.377-383.

alternate forms. Each examiner saw one accomplice who acted warm and another who acted cold. Each accomplice had an equal number of warm and cold roles in the experiment. The interpretations of the protocols were themselves analyzed for the numbers of "positive", "negative", and "other" statements. Two judges agreed eighty-one per cent in their classification, while the mean and median agreement per protocol was seventy-seven per cent. The results of the experiment indicated that of two "equivalent" protocols an examiner made more positive statements about the one from the "warm" subject than about the one from the "cold" subject. Also, when subjects were warm to the examiner they were perceived more positively than when they were cold.

In an attempt to predict psychiatric patients' work adjustment, Stotsky and Weinberg⁴⁷ utilized eighty-one stems, sixty-nine of those being related to nine ego-strength variables and twelve being neutral. The nine ego-strength variables were: reactions to situations of difficulty; need achievement; specificity of goals; reactions to failure; self-reliance; persistence on the job in the face of boredom, distractions, temptations, and the

⁴⁷ Bernard A. Stotsky and Henry Weinberg, "The Prediction of the Psychiatric Patient's Work Adjustment", in the Journal of Counseling Psychology, Vol.3, No.1, issue of Spring 1956, p.3-7.

prospect of hard work; reaction to superiors; reactions to peers; and, reaction to subordinates. Items were arranged so that stems relating to any one dimension were not clustered together. The test was given to eighty patients in the Manual Arts and Educational Therapy subsections of a Veterans Administration hospital. In the group were seventy-eight diagnosed schizophrenics, one sociopathic personality disturbance and one depressive reaction. Each response was scored plus 1 for positive, minus 1 for negative, or zero for neutral, and these scores were then summed algebraically for variable and total scores. With three judges the experimenters found the interrater agreement to be greater than ninety per cent on four-hundred individual completions.

The test was validated against the following criteria:

1. A 23-item rating scale completed by therapists, to give a high and low group in terms of performance ratings.
2. The criterion of adjustment six months after the test was given.

With the first criterion, eight of the variables gave indications that were significant at the .05 level of confidence, and five variables gave indications that were significant at the .01 level. The "self-reliance" variable was not significantly related to work performance ratings. With the second criterion (predictive validity), two of the variables were significantly related to the criterion of

adjustment at the .05 level of confidence; and, seven were significant at the .01 level. In the second case, "self-reliance" was the variable that best differentiated those patients with positive outcomes from those with negative outcomes. The authors explained

. . . . self-reliance items may be measuring the striving and determination of patients to ascertain their independence and break their identification with the hospital community. If this is so, self-reliance would be of much greater significance in determining outcome, which involves progress toward or actual preparation from the hospital, than in assessing adjustment to work activities, since the latter places greater emphasis on interpersonal factors.⁴⁸

In a further study, Stotsky⁴⁹ used the same test with seventy-eight "relatively nonregressed" schizophrenics and thirty-two male normals, matched as groups by age and education. The latter were either Veterans Administration employees or students in an evening course. The schizophrenics were divided into two sub-groups of thirty-nine each, one being with a poor outcome of treatment, the other with a positive outcome in treatment. The tests were scored "by one whose reliability had been previously established". It was reported that the normals obtained a significantly (at

⁴⁸ Ibid., p. 6.

⁴⁹ Bernard A. Stotsky, "Comparison of Normals and Schizophrenics on a Work-Oriented Projective Technique", in the Journal of Clinical Psychology, Vol.13, No.4, issue of October 1957, p.406-408.

the .01 level of confidence) greater number of high scores than did the schizophrenics for six of the nine variables. Another variable showed a difference between groups that was significant at the .05 level of confidence. Two variables, "specificity of goal" and "reaction to failure", showed no significant differences. All the significant differences were in the expected direction.

The normals had significantly (at the .01 level) greater number of high scores on eight of the nine variables and on the over-all score than did the schizophrenic subgroup with a poor outcome of treatment. The normals were significantly higher on two of the variables and on the over-all score than the schizophrenics with a positive outcome in the rehabilitation treatment.

This chapter has presented an overview of the development of the sentence completion technique from its use by Payne in 1928 to 1947, when the Rotter Incomplete Sentences were evolved. The second half of the chapter dealt mainly with the use of the technique from 1948 to 1957. In the next chapter is presented a review of those studies which dealt specifically with the types of items used in this technique, with particular reference to the use of self-referent and other-referent items.

CHAPTER III

REVIEW OF LITERATURE RELATED TO TYPES OF ITEMS

In the main, three types of incomplete sentences have been used in the make-up of sentence completion tests; self-referent, other-referent and non-personal items. Some authors have considered all three types to be projective items; others hold that one or the other alone is projective. The following survey reviews the studies and writings related to those points of view and indicates the lack of conclusive evidence that can be gained from these reports.

1.- In General

At the twenty-seventh annual convention of the Western Psychological Association in 1947, Lehner¹ read a paper on the "Projection of Men and Women to Items Referring to the Same and Opposite Sex on a Sentence Completion Test".

A sentence completion test was constructed to contain stimulus material that would "force" projections to items referring to either members of the same or of the opposite sex. Of the hundred items, fifty were from the male and fifty from the female viewpoint. The first fifty

¹ G.F.J. Lehner, "Projections of Men and Women to Items Referring to the Same and Opposite Sex on a Sentence Completion Test", an abstract in the American Psychologist, Vol.2, No.10, issue of October 1947, p.407.

were matched as far as possible with those in the second set of fifty items. The test was administered to one hundred male and one hundred female subjects. Half of the males and the same number of the females had no siblings. Of the remainder, fifty males had one or more sister and fifty females had one or more brother. The responses of the groups to the "male" and "female" items were compared. The author's results and conclusions were not published with the abstract of his paper.

Meltzoff² investigated some of the basic assumptions of projective techniques of the sentence completion type by testing two main experimental hypotheses:

HYPOTHESIS I: Other things being equal, the tone and neutrality of the responses to a sentence completion test are direct functions of the S's mental set, as determined essentially by test instructions.³

HYPOTHESIS II: Other things being equal, response tone is a direct function of the tone of the stimulus.⁴

2 Julian Meltzoff, "The Effect of Mental Set and Item Structure upon Response to a Projective Test", in the Journal of Abnormal and Social Psychology, Vol.46, No.2, Issue of April 1951, p.177-189.

3 Ibid., p.179.

4 Ibid., p.180.

A sentence completion test containing an equal proportion of three types of items ("Positive", "negative", and "neutral") was administered under four different conditions to 120 college students. Testing conditions were such as to provide two mental sets that favoured distortion in responses by asking the subjects to act the role of well-adjusted and emotionally disturbed individuals, a condition in which self-esteem was threatened, and a condition under which self-esteem threat was reduced by assurance of anonymity. The responses were classified by clinical psychologists as positive or suggestive of good adjustment, negative or suggestive of poor adjustment, and neutral or evasive.

Meltzoff concluded that the results of the study supported both of his main hypotheses. In reference to the first hypothesis, he stated:

- a) Subjects are able to manipulate responses on a projective test of this type so as to create the impression of either good or poor adjustment.
- b) When self-esteem is threatened by the test situation, the subjects respond practically the same manner as those who are instructed to act well-adjusted with the exception that they give more evasive responses and take longer time to respond.
- c) Subjects working anonymously, and without having self-esteem threatened by the test situation, respond most rapidly and give significantly more maladjusted responses along with fewer well-adjusted and evasive responses than those who are self-esteem threatened.⁵

⁵ Ibid., p.188-189.

With reference to the second hypothesis, he concluded:

- a) Positive stimuli tend to elicit responses that are suggestive of good adjustment.
- b) Negative stimuli tend to elicit responses that are suggestive of poor adjustment, and allow least freedom of response.
- c) Neutral stimuli do not directly affect the tone of the responses and allow most freedom of response.⁶

Two other authors, Meyer and Tolman⁷ compared attitudes to the "mother" and to the "father", as given by twenty patients in response to eleven items of the Forer Sentence Completion Test, the Thematic Apperception Test, and in therapy sessions. Of the eleven sentence completions used, three were self-referent and eight were of the other-referent type. The authors concluded that the attitudes elicited on one measure furnished no basis of prediction of those found on either of the other measures. It would be of considerable interest to repeat this study using a more adequate experimental design.

6 Ibid., p.189

7 Mortimer M. Meyer and Ruth S. Tolman, "Parental Figures in Sentence Completion Test, in TAT and in Therapeutic Interviews", in the Journal of Consulting Psychology, Vol.19, No.3, issue of June 1955, p.170.

Osterweil and Fiske⁸ studied the variability in responses and relative response frequency to a sentence completion test over a two to three week interval. The major results reported were:

1. The content of the majority of responses was changed on retest.
2. The individual differences in the number of responses changed showed reasonable reliability.
3. Differences between sentence stems in the proportion of changes responses were consistent from group to group.
4. For each item responses given by only one person changed more frequently than those given by more than one person.
5. Positive correlations were found between the number of unique responses made by a person and the number of responses that person changed. Similarly, negative correlations were found between the number of populars and the number of changed responses per subject.

The authors concluded that change of responses on a sentence completion test can be attributed in part to

⁸ Jerry Osterweil and Donald W. Fiske, "Intra-Individual Variability in Sentence Completion Responses", in the Journal of Abnormal and Social Psychology, Vol.52, No.2, issue of March 1956, p.195-199.

factors associated with individuals and to factors associated with items.

2.- Supporting the Use of Self-Referent Items

Self-referent items were used in preference to other-referent items by Rohde⁹ as the former were found by her to give greater productivity of significant responses. She carried out a study with some other-referent items in preparation for the 1940 Sentence Completion Test "but concluded that these items did not yield data equal in significance to those yielded by first-person items".

Sacks¹⁰ reported on his research conducted with a view to investigating the validity of the assumption that a subject revealed more about himself when he talked about other people or about impersonal unstructured objects than when he talked about himself. For this study, Sacks constructed two forms of a sentence completion test: Form A, in which all the items were of the self-referent type; and Form B, in which all items were of the other-referent type.

⁹ Amanda R. Rohde, The Sentence Completion Method, Its Diagnostic and Clinical Application to Mental Disorders, New York, Ronald Press, 1957, p.52 with footnote.

¹⁰ J.M. Sacks, "The Relative Effect upon Projective Responses of Stimuli Referring to the Subject and of Stimuli Referring to Other Persons", in the Journal of Consulting Psychology, Vol.13, No.1, issue of February 1949, p.12-20.

To obtain the items for the test, twenty psychologists were asked to submit three incomplete sentences for each of fifteen categories. Sacks added other items to form a pool of 280 items, from fourteen to twenty-six items per category. These were then submitted to the same twenty psychologists who were requested to select the four items in each category which they thought best suited to elicit the subject's attitudes in that category. In each of the categories, the four incomplete sentences most frequently chosen became the test items. The sixty items were then assembled so that each category was represented once in each set of fifteen items. In addition, each item was worded with a pronoun of the first person singular in it. Form B was made by changing the pronoun to one in the third person or to a proper noun. For example, "If I were in charge" became "If Bob were in charge", "If my father would only" became "If his father would only".

The two forms of the test were administered to ninety-four male and four female patients of the Veteran's Administration New York Regional Office, Mental Hygiene Clinic. These patients ranged in age from nineteen to fifty-one years, the median age being twenty-seven and a half. The group consisted of fifty-nine neurotics, twenty-four psychotics, thirteen patients with character disorders, and four with narcoleptic or organic disorders. These subjects were numbered consecutively and the odd numbers given

Form A of the test first then Form B. The even numbers were given these in the reverse order.

The completions for each form of the test were typed on separate rating sheets. These were then analyzed in terms of four major areas of personality:

1. Family Attitudes -- toward father, toward mother, toward family unit.
2. Sex Attitudes -- toward women, toward heterosexual relationships.
3. Interpersonal Attitudes -- toward friends and acquaintances, toward superiors at work or school, toward people supervised, toward colleagues at work or school.
4. Self Attitudes -- fears, guilt feelings, attitudes toward own abilities, attitude toward past, attitudes toward future, and goals.

Three psychologists independently rated the responses in each category. A rating of 2 was given to a response thought to be indicative of a severe disturbance; a rating of 1 was given to responses thought to be indicative of a mild disturbance; all others received a 0 rating. These ratings were compared with psychiatrists' clinical impressions of the same individuals on the same variables.

Sacks reported that the self-referent type of item yielded the greatest degree of similarity between the psychologists' and the psychiatrists' judgments in several

categories; attitudes toward mother, father, family unit, heterosexual relationships, superiors, colleagues, future, and goals. He suggested that further investigation was needed on the use of an other-referent form of the test in which only pronouns were used, since proper names tended to arouse associations with specific persons.

Sacks also found a larger proportion of responses on the other-referent form expressed negative feelings toward the father and toward superiors, but he made no attempt to suggest any explanation for this.

Rotter and Rafferty described the items of the Rotter Incomplete Sentences Blank as follows:

The items may be characterized in general, as being short and unstructured -- as in the sentences beginning "I . . . " and "Sometimes . . . " -- and as being either neutral or referring to the first person. Third person stimuli, such as "He always worried . . .", or "Mary felt that . . . ", are not used in the ISB. The authors' experience indicated that such items tended to produce stereotypes and to be literally referred to other people.¹¹

Cromwell and Lundy¹² reported on research carried out to investigate the productivity of stems for clinical

¹¹ Julian B. Rotter and Janet E. Rafferty, Manual, the Rotter Incomplete Sentences Blank, College Form, New York, Psychological Corporation, 1950, p.6; and, "The Rotter Incomplete Sentences Blank", in Arthur Weider, Edit., Contributions toward Medical Psychology, Theory and Psychodiagnostic Methods, Vol.2, New York, Ronald Press, 1953.

¹² Rue L. Cromwell and Richard M. Lundy, "Productivity of Clinical Hypotheses on a Sentence Completion Test", in the Journal of Consulting Psychology, Vol.18, No.6, issue of December 1954, p.421-424.

hypotheses. They used sixty-five items, forty of these were from the Rotter Incomplete Sentences Blank, Adult Form, and the remaining items were constructed with a view to securing further information as to hospital ward adjustment and patient identification. The authors were concerned also with the utility of the other-referent type of item in which a personal pronoun was used.

The test was administered to sixty neuropsychiatric admissions at a Veterans' Administration hospital. The order of the stems and two different sets of instructions were alternated from subject to subject. Thirty-nine clinical psychologists in various hospitals made personality inferences from the protocols, and listed the completed sentences used in making the inferences. The authors' findings were:

1. that individual differences in stem productivity were significantly consistent from the first to the second half of the group of subjects;

2. that forty-five of the stems could be useful for psychological evaluation in neuropsychiatric hospitals;

3. that the more productive stems were of the self-referent type, referred to the present or the future, and to the 'emotional' aspects (e.g. hostilities, worries, fears, and troubles) of the subjects;

4. that the different instructions did not affect test productivity.

Izard, et al,¹³ in constructing and validating a multiple-choice sentence completion test, wrote the stems in the self-referent form whenever the item content permitted. The authors concluded from their study:

The present results, especially with respect to its reliability and its validity against peer and self-ratings of social adequacy, appear to warrant further development of the instrument.¹⁴

3.- Supporting the Use of Other-Referent Items

Stein¹⁵ reported on the sentence completion test constructed under the direction of Murray for the Office of Strategic Service selection programme. The test was in two parts, each of fifty items. These items were selected on the basis of their contributing to any of twelve areas considered important for personality evaluation. These were: Family; The Past; Drives; Inner States; Goals; Cathexes; Energy; Reaction to Frustration and Failure; Time-Perspective; Optimism-Pessimism; Reactions to Others; Reactions of Others to the Candidate. The items were selected for the test also on the basis of giving a high index of individuality,

13 Carroll E. Izard, Nathan Rosenberg, John T. Bair and Clinto Maag, Construction and Validation of a Multiple-choice Sentence Completion Test, an Interim Report, Project Number NM001 077.01.02, Report No.2, Pensacola, U.S. Naval School of Aviation Medicine, 1953, 9-x p.

14 Ibid., p.8.

15 M.I. Stein, "Use of a Sentence Completion Test for the Diagnosis of Personality", in the Journal of Clinical Psychology, Vol.3, No.1, issue of January 1947, p.47-56.

that is, less than fifty per cent of the responses to an item were similar. Again, two types of items, self-referent and other-referent, were mixed in random order. From the test as a whole the clinician wrote a personality description of the subject. Stein considered that "It offers the clinician clues as to critical areas in the personality and some knowledge of the dynamics of behaviour involved in cases under observation".¹⁶ The more projective items, according to the author, were those in which the proper name of some masculine person or the third personal pronoun were employed (other-referent items).

Symonds¹⁷ reported on the same test. He was more objective than Stein and described the attempted validation of the items. Twenty-five candidates for the Office of Strategic Services¹⁸ with high over-all ratings and high emotional stability ratings on their assessments were compared with a group of twenty-five candidates who had low ratings in these categories. The responses on the sentence completion

16 Ibid., p.56

17 P.M. Symonds, "The Sentence Completion Test as a Projective Technique", in the Journal of Abnormal and Social Psychology, Vol. 42, No. 3, issue of July 1947, p.320-329.

18 U.S. Office of Strategic Service Assessment Staff, Assessment of Men, New York, Rinehart, p.71-75; and Henry A. Murray and Donald W. McKinnon, "Assessment of OSS Personnel" in the Journal of Consulting Psychology, Vol. 10, No. 2, issue of March-April 1946, p.76-80.

test were categories: persistence; striving for success; feelings of inferiority, doubt and worry; depression, discouragement following failure; high standards; emotional stability in stressful situation. Symonds reported that the responses did not differentiate between groups and concluded:

The interpretation of sentence completions as projections is doubtful, questionable, hypothetical, conjectural, and supposititious, and cannot be the substitute of a complex analysis. The sentence completion test is not a safe guide at the present time for prediction or assessment.¹⁹

However, Symonds himself admitted that the criterion used for this validation study was somewhat inadequate.

Getzels²⁰ compared the results of a group of twenty-five clinically judged "well-adjusted" veterans and forty clinically judged "maladjusted" veterans on an other-referent sentence completion test with a self-referent, multiple choice, sentence completion test. Twenty of the items on the latter matched twenty of the forty items of the former.

On the other-referent form the maladjusted group showed considerably more maladjustment than did the well-adjusted group. However, on the multiple choice form there were no significant differences between the two groups. Again, the author reports more inconsistency in the responses to the

19 P.M. Symonds, op.cit., p.328.

20 J.W. Getzels, The Assessment of Personality and Prejudice by the Methods of Paired Direct Projective Questionnaires, unpublished Doctoral thesis presented to the Department of Social Relations of Harvard University, Cambridge, 1951.

matched other-referent and self-referent items for the maladjusted group. Getzels stated that the items were designed to relate to six categories of vocation and personal pertinence and included a number of neurotic symptoms. He concluded that since the maladjusted group were anxious to make a favourable impression to receive certain financial grants, it was suspected that the performance on the self-referent measure was artificial and indicated defensiveness.

It would appear that there was more than one variable being manipulated in the study. In the first place the twenty items were changed from the other-referent to the self-referent type. Secondly, the form of the responses was changed from a projective association completion to a direct, selective response.

Wallon and Webb²¹ reported on experiments in which the Rosenzweig Picture-Frustration test and a sentence completion test were modified into multiple-choice forms. Those tests were given in three ways: the projective tests alone, the multiple-choice forms alone, and the projective tests given and later the subjects required to match their responses to the multiple-choice forms. The "objectification" resulted in a marked increase in "socially acceptable" responses. However, the joint matching of projective to objective forms

21 Edward J. Wallon and Wilse B. Webb, The Effect of Varying Degrees of Projection on Test Scores, Project No. NM 001 108 100, Report No. 12, Pensacola (Florida), U.S. Naval School of Aviation Medicine, 1956, 11-20p.

approximated the projective alone. Applying these results to Getzels study, it is questionable if his findings resulted from a change from the other-referent to the self-referent form as he concluded.

In the second experiment, reported by Getzels,²² a 59-item sentence completion test, composed of other-referent stimulus items, was given to three classes in psychology at a women's college. Two forms of a self-referent sentence completion test made up of forty items each that matched forty of the fifty-nine items given initially were then administered to the three classes. Test A was in multiple choice form and Test B was made up of open-ended stimuli. The first class was given Test A and the individuals were asked to sign their names for the college records office. The second class also received Form A but the students were instructed not to sign their names as the data were for research purposes. The third class was given Form B and its various members were given the same instructions as the first class.

The tests were scored for the category "prejudice". In all three groups the other-referent scores were higher than the scores obtained on either of the self-referent forms. The least discrepancy between self-referent and other-referent scores

22 J.W. Getzels, op. cit.

appeared in the group which had been guaranteed anonymity. The distribution of the prejudice scores on the self-referent forms tended to approximate a J-curve, whilst the scores on the other-referent test were normally distributed.

From those data Getzels concluded that self-referent items on the sentence completion test are not truly projective stimuli in that subjects responding to socially conflicted objects of inquiry may express their reactions to conform with social expectation.

It is interesting to note that Getzels used highly structured²³ stimulus items in all the test forms. This in itself may have influenced the results found, for as Bellak²⁴ has noted, projection is inversely proportional to the clearness of the stimulus. In comparison, the tests used in clinical practice more often have items that are relatively unstructured.

23 Hutt differentiates between unstructured, partially structured and structured test as follows: Unstructured personality tests "are psychological experiments designed to obtain penetrating insight into the personality structure of the subject. Neither the stimulus nor the response is conventionalized. There is no restriction of the subject's use of the stimulus or his responses to it".

Partially structured personality tests are "those measures in which the test stimulus remains relatively structured or conventionalized, but the subject may respond in a relatively free or individual manner to this stimulus".

Structured personality tests are "those tests in which the test material consists of conventionally, culturally crystallized questions to which the subject must respond in one of a very few fixed ways". See Max L. Hutt, "The Use of Projective Methods of Personality Measurement in Army Medical Installations" in the Journal of Clinical Psychology, Vol. 1, No. 2, issue of April 1945, p. 135-136.

24 Leo Ballak, "The Concept of Projection, an Experimental Study of the Concept", in Psychiatry, Vol. 7, issue of 1944,

In their study of the sources of reference in response to sentence completion items, Hanfmann and Getzels²⁵ found that many responses to other-referent items were determined by unconscious self-reference.

Kimball²⁶ gave fifty other-referent sentence completions, in which masculine personal names and some personal pronouns were used, to seventeen male, adolescent, scholastic under-achievers. Her purpose was to investigate the relevance of school achievement to father-son relationship and school achievement to aggression toward living things as a source of guilt and anxiety. As a control group Kimball used one-hundred subjects selected at random from the total school population in a preparatory school.

There was perfect agreement found between the overall ratings of the experimental group's tests made by the author and other judge. There was ninety-six per cent agreement between the two rate on single items.

In the results as reported, the author found that fifty-nine per cent of the experimental group were judged to have predominantly negative relationships with their fathers as

25 Eugenia Hanfmann and J.W. Getzels, "Studies of the Sentence Completion Test", in the Journal of Projective Techniques, Vol. 17, No. 3, issue of September 1953, p.280-294.

26 Barbara Kimball, "The Sentence-Completion Technique in a Study of Scholastic Underachievement", in the Journal of Consulting Psychology, Vol. 16, No. 5, issue of October 1952, p. 353-358.

compared to twenty-nine per cent of the control group. Again, forty-seven per cent of the underachievers introduced aggression towards living things in their completions as compared with thirteen per cent of the control group.

Getzels and Walsh²⁷ used the "Method of Paired Direct and Projective Questionnaires (PDPQ)" in a study of attitude structure and socialization. In this technique, sentence completion stems were paired, being identical except that the first pronoun was used in one and a proper name in the other. The authors made the assumptions that

In responding to the direct instrument, the subject is fully aware that what he says refers to himself, and that a value judgment within some particular frame of reference of which he is also aware may be made on the basis of this responses

In this study, the position is taken that the response to a direct question represents the level of behaviour at which the individual permits society to look at him - it is the expressed reaction of the preceding conceptualization of the question - answer process.

In responding to the projective instrument, the subject is unaware that what he says refers to himself, and that a value judgment within a particular frame of reference may be made on the basis of his responses

The basic assumption is that when a subject is pressed to attribute behaviour to others, and he perceives the situation as devoid of personal reference, he tends to reveal levels of behaviour he might otherwise be disposed to distort or censor.²⁸

27 J.W. Getzels and J.J. Walsh, "The Method of Paired Direct and Projective Questionnaires in the Study of Attitude Structure and Socialization", in Psychological Monographs, Vol. 72, No. 1 (Whole No. 454), issue of 1958, 34 p.

28 Ibid., p.3.

Further, to disguise the nature of the "projective questionnaire" it was presented as a verbal speed test.

One criticism might be voiced, at this point, of the term "questionnaire". The tests used were not questionnaires in the sense in which the term is normally used in psychological and sociological measurement. Further, the items themselves in most cases were highly structured.

The authors after describing the rationale, construction, scoring, and interscorer reliability of the technique, presented the results of two experimental studies. The first was concerned with "the nature of the relationship between overt and covert levels of attitude structure as a function of differential objects of inquiry"; the second, "with the nature of the relationships between overt and covert levels of attitude structure as a function of differential patterns of socialization represented by the age, sex, sibling position, and social class characteristics of respondents".²⁹

The authors found that normals have more socially desirable or positive responses under direct self-reference. In the second experiment, this tendency was found to increase with age between eight and thirteen years. The authors suggested that as evidence of the socialization process. Also, the extent to which boys and girls, with age and IQ constant, tend to inhibit their socially unacceptable personal hypotheses (i.e. their true

²⁹ Ibid., p.2.

social attitudes) was found to be clearly a function of the socioeconomic class membership of those children.

In a study of the "authoritarian personality" by means of a sentence completion technique, Dorris, Levison and Hanfmann³⁰ were able to confirm twelve of sixteen hypotheses about the authoritarian personality make-up by comparing ten subjects high in authoritarianism and eleven subjects rated low in authoritarianism by means of three authoritarianism scales.

As part of their experiment, these authors had each subject state whether each completion was true of himself or not subsequent to taking the sentence completion test. It was found that other-referent items elicited significantly more denials of self-reference than did completions to the self-referent type of item. It was concluded that "This finding is consistent with the frequently made assumption that third-person items facilitate the expression (projection) of tendencies that the subject does not recognize in himself".³¹

Davids³² reported a study designed to compare measures of effectiveness of adjustment derived from his hypothesized three methods of personality assessment direct, indirect and projective. Two experiments were carried out. In both,

30 Ronald J. Dorris, Daniel J. Levison and Eugenia Hanfmann, "Authoritarian Personality Studied by a New Variation of the Sentence Completion Technique", in the Journal of Abnormal and Social Psychology, Vol. 49, No. 1, issue of January 1954, p.99-108.

31 Ibid., p.99.

32 Anthony Davids, "Comparison of Three Methods of Personality Assessment, Direct, Indirect and Projective", in the Journal of Personality, Vol. 23, No. 4, issue of June 1955, p. 423-440.

normal subjects were tested by the same five experimental measures; a psychosomatic inventory, a self-rating technique, a personal interview, a sentence-completion test, and a word-association technique. The experimental variable was the motivational conditions operating with each group. The first group consisted of twenty male undergraduates who voluntarily participated for research reasons. With this group anonymity was assured. The second group was made up of twenty-three male undergraduates who had applied for "non-existent" high paying jobs as a research assistant in the psychology department. These were told that the test battery would be used to determine several of the applicants who appeared to be free from personal problems.

The study substantiated Allport's assertion that the personality test performance by the normal person will be consistent in direct and projective measures. There was a significant degree of agreement among rank orders on maladjustment scores for each of the five measures with both groups. However, with the second group the intercorrelations among different measures and the coefficient of concordance were considerably lower in size than in the first experiment. When the two groups were compared it was found that the subjects in the second group evidenced less maladjustment. The differences were of acceptable significance on only two of the measures, however. The author concluded that performance on personality measures is a function of many variables, one of which is motivational conditions.

Davids and Pildner³³ reported on a more extensive study of the relative performance on direct and projective personality assessment measures by subjects under different motivational conditions. A battery of nine personality assessment tests was assembled to include four projective measures: a word association test, a sentence completion test, the Azzageddi Test, and the Thematic Apperception Test, and five direct instruments: a self-rating scale, the Happiness Scale, the Affect Questionnaire, the Psycho-Somatic Inventory of McFarland and Seitz, the Taylor Manifest Anxiety Scale. In addition, a clinical evaluation of each subject was made on the basis of an initial interview by a psychologist prior to the testing program.

Two groups of subjects were studied. The control group consisted of twenty male undergraduate volunteers who were paid for their time, guaranteed anonymity and were encouraged to be honest and uninhibited in the cause of science. The experimental group consisted of twenty-three male undergraduates, not majoring in Social Relations or Psychology, who had applied for employment as a research assistant at the Harvard Psychological Clinic.

33 Anthony Davids and Henry Pildner, "Comparison of Direct and Projective Methods of Personality Assessment under Different Conditions of Motivation", in Psychological Monographs, Vol. 72, No. 11, (Whole No. 464) issue of 1958, p.30.

As the Authors pointed out:

It was expected that the experimental subjects would attempt to distort the personality measures in an effort to appear "nature and well adjusted". Because of the nature of the projective test material, it was predicted that the experimental group would be unable to alter significantly their performance on the projective measures so as to appear exceptionally well adjusted or in this case, nonalienated. However, since the direct instruments were far more transparent, it was expected that they would be quite susceptible to distortion. Since the control group subjects were motivated to respond honestly to these assessment measures, the experimenters predicted that the experimental group would show significantly less alienation on the direct measures and would not differ from the control group on projective test alienation scores. Also, it was predicted that the experimental group would perform less consistently on direct and projective tests and, therefore, in comparison with the control group, would show less concordance among rank orders based on the results of the various measures.³⁴

The results demonstrated that distortion is likely to take place on direct personality measures as the result of the motivational situation. Likewise, projective methods can be used effectively to control assessment deception.

It is important to note that of the hundred sentence completion items, half were in the first person (self-referent) and the other half in the third person. The authors do not state if these were matched items. These were mixed in random order. In the experiment reported only third-person (other-referent) items were used in computing alienation scores for the sentence completion test. Considering the hypothesis that self-referent items are not projective stimuli, the mean alienation

³⁴ Ibid., p.27.

scores on the self-referent items for the control and experimental groups were compared. The difference between the mean scores was significant at the 5 per cent level of probability. The experimental group showed less alienation. Rank-order correlations between alienation scores on the self-referent and other-referent items were computed for each group. A correlation of .83 was obtained for the control group, whereas in the experimental group the correlation was .44. The authors pointed out that although the latter correlation was statistically significant, it left approximately eighty per cent of the variance unexplained. They concluded that self-referent items were not as efficient as other-referent stems in eliciting unconscious or inhibited material and the responses to them could be distorted.

4.- Supporting the Use of Both Types of Items

Forer³⁵ described a 100-item test in which he attempted to achieve the maximum structuring of the stems. He felt that minimally structured stimulus items left the underlying psychological processes in doubt and that in such cases an interpreter would lack sufficient information regarding the stimulus situations to determine what the content of the responses meant.

³⁵ Bertram R. Forer, "A Structured Sentence Completion Test", in the Journal of Projective Techniques, Vo.14, No. 1, issue of March 1950, p.15-30.

Four forms³⁶ of his test have been published (Form M for males, Form B for boys, Form F for females and Form G for girls). Forms M and F are virtually identical except for changes in pronouns and positions of the items. The same applies to Forms B and G. For scoring the Forer Structured Sentence Completion Test Record Form is used with all four. In the Manual³⁷ Forer stated that he framed the items "in three basic forms: first person singular, third person singular, and third person plural".³⁸ The author made the assumption that the other-referent forms would reveal some material which might be threatening to the individual in the self-referent type.

Forer and Tolman³⁹ reported on a study conducted with FSSCT. A group of clinical psychologists rated the items of that test in terms of the degree of potential clinical value. Also indicated was the degree of confidence of their own

36 Bertram R. Forer, The Forer Structured Sentence Completion Test (FSSCT), Forms M, B, F and G, Los Angeles, Western Psychological Services, 1957.

37 Bertram R. Forer, Manual, The Forer Structured Sentence Completion Test (FSSCT), Los Angeles, Western Psychological Services, 1957, 15 p.

38 Ibid., p.5.

39 Bertram R. Forer and Ruth S. Tolman, "Some Characteristics of Clinical Judgment", in the Journal of Consulting Psychology, Vol. 16, No. 5, issue of October 1952, p. 347-352.

judgments. These ratings were then studied in relation to

1. the type of item (self-referent or other-referent),
2. the area of content, and
3. the degree of certainty or confidence expressed in the rating.

The authors concluded that

1. No significant differences in value judgments were found between self-referent and other-referent items;
2. Items dealing with causes of feelings and with reactions to interpersonal relationships were judged to be more clinically valuable than items dealing with attitudes toward various interpersonal figures;
3. Confidence ratings were found to be unrelated to objective features of the test material, except that increased experience in the rating procedure produced a significant increase in confidence of one's judgments;
4. The differential confidence ratings which individual clinicians made with respect to their own clinical judgments appeared to be consistent, systematic, personal characteristics;
5. Most clinicians felt more confident when in disagreement rather than in agreement with the group, especially where the general confidence levels of those clinicians were high;
6. Confidence was associated more with extreme than with moderate judgments, especially in those clinicians with a generally high confidence level;

7. High confidence appeared to be associated with positive rather than negative judgments of the clinical usefulness of items.

Trites⁴⁰ reported on a factor analysis of the inter-correlations of seventy-four of the eighty-eight items of the School of Aviation Medicine Sentence Completion Test.⁴¹

Forty-four of those items used were of the self-referent type. Of the four factors he found, the Self-Centred Anxiety and the Narcissistic (or Self-Enhanced) factors seemed best defined by items containing a self-reference. The results seem to support a theory that self-referent items might define some areas of personality, whilst other types of items might define other areas.

This chapter has reviewed the literature pertaining to the use of various types of sentence completion items. In particular, it has dealt with studies on the use of the

40 David K. Trites, Basic Dimensions and Predictive Validity of a Projective-Type Sentence Completion Test, unpublished Doctoral Thesis presented to the University of Texas, Austin, 1953; Psychiatric Screening of Flying Personnel, Evaluation of Assumption Underlying Interpretation of Sentence Completion Tests, Report No. 55-33, Randolph Field, (Texas) U.S.A.F. School of Aviation Medicine, 1955, 9 p; and "Evaluation of Assumptions Underlying Interpretation of Sentence Completion Tests", in the Journal of Consulting Psychology, Vol. 20, No. 1, issue of February 1956, p. 8.

41 The construction and validation of that test was described by D.K. Trites, W.H. Holtzman, R. Templeton and S.B. Sells, Psychiatric Screening of Flying Personnel, Research on the SAM Sentence Completion Test, Project No. 21-0202-0007, Report No. 3, Randolph Field, (Texas) U.S.A.F. School of Aviation Medicine, 1953, 23 p.

self-referent and the other-referent forms of stimulus items. In all of the reports surveyed conclusive evidence as to the value of one or the other form of an item is lacking. It was with the need for more evidence in mind that the experiment recorded in the following chapter was organized and performed.

CHAPTER IV

THE PRESENT STUDY

Two groups of adult males differing in socio-emotional adjustment were used in an experiment designed to test hypotheses put forward with regard to the assumptions underlying self-referent and other-referent sentence completion items. The following gives a statement of the problem, the rationale on which the study was based, the general hypotheses proposed, the manner in which the subjects were obtained, and the procedure used.

1.- Statement of the Problem

In Chapter I, the concept of projection as applied in the term "projective test" was discussed with a view to clarifying the views of many authors, as to what constituted projection in a test. Whilst authors generally concurred, they differed by the application of certain specified theoretical assumptions to the technique. It was seen that when the term was applied to the sentence completion test, it was through the use of introspection and by the application of logical analysis that authors concluded the latter device to be one of the many projective methods.

In the review of the literature dealing with the sentence completion technique, which was presented in Chapters II and III, it was noted that the theoretical assumptions grew up after the use of the method. However, it was not until 1949

that a research designed to test one of those theoretical assumptions was reported. That study was conducted by Sacks with a view to investigating the validity of the assumption that a subject revealed more about himself when he talked about other people or about impersonal unstructured objects than when he talked about himself. Although Sacks did not accept his findings as conclusive, many authors since have quoted them as if they were. In Chapter III, it was noted that studies since that time both clinical and experimental, have failed to contribute to the over-all agreement as to the projective or non-projective nature of the self-referent and other-referent stimulus items in a sentence completion test.

The problem revolves around three specific questions:

1. Do responses of individuals to self-referent and other-referent stimuli in a sentence completion test differ significantly in terms of the amount of unfavourable attitudes, disturbances, and problems revealed?

2. Is the self-referent type of stimulus item more revealing of maladjustment than the other-referent form of the same item or vice versa?

3. Are both types of items "projective" techniques?

It was considered that in the examination of the third question, answers to the other two should be obtained. What was required was an objective means of assessing a technique to judge whether it be a direct or a projective one.

A. The Rationale Upon Which the Study Was Based

Allport¹ has suggested that direct and projective measures of assessment will reveal the same information about the motives and attitudes of the normal, well-adjusted individual. However, with the neurotic, maladjusted person, projective techniques will elicit repressed material which is contrary to that elicited on direct measures of personality. The maladjusted person unconsciously or consciously distorts his responses to direct measures toward the socially accepted norms in order to maintain and to enhance his ego. It follows, that whilst responses to direct or non-projective instruments may be faked or distorted, those responses to projective devices are insusceptible to distortion because of the ambiguity of the stimulus material.

Underlying that assertion is the hypothetical construct of behavioural levels of reaction. These levels have been variously named by different authors, as for example, "private-public",² "unconscious-conscious",³

1 G.W. Allport, "The Trend in Motivational Theory", in the American Journal of Orthopsychiatry, Vol. 23, No.1, issue of 1953, p.107-119.

2 S. Rosenzweig "Levels of Behaviour in Psychodiagnosis with Special Reference to the Picture Frustration Study", in the American Journal of Orthopsychiatry, Vol. 8, issue of 1938, p.622-626.

3 Sigmund Freud, A General Introduction to Psycho-Analysis, New York, Liveright, 1935; translated and edited by Joan Riviere. (First German edition 1917).

"genotype-phenotype",⁴ "covert-overt",⁵ "latent-manifest",⁶ and so forth. What Allport contended was that direct techniques of personality measurement tap the overt level of behaviour, whereas projective measures elicit information pertaining to the covert level. In the normal, well-adjusted individual those levels of behaviour coincide, whereas in the neurotic those levels diverge.

That formed a much need basis for the operational definitions of "direct" and "projective" with reference to the measurement of personality. Particularly as the studies reported by Laslett and Bennett,⁷ Spencer,⁸ Combs,⁹

4 D.W. MacKinnon, "Psychodiagnosis in Clinical Practice and Personality Theory", in the Journal of Abnormal and Social Psychology, Vol. 44, No. 1, issue of January 1949, p. 7-13.

5 D.T. Campbell, "The Indirect Assessment of Attitudes", in the Psychological Bulletin, Vol. 47, No. 1, issue of January 1950, p. 15-38.

6 E. Hanfmann and J.W. Getzels "Studies of the Sentence Completion Test", in the Journal of Projective Techniques, Vol. 17, No. 3, issue of September 1953, p. 280-294.

7 H.R. Laslett and E.A. Bennett, "A Comparison of Scores on Two Measures of Personality", in the Journal of Abnormal and Social Psychology, Vol. 28, No. 4, issue of January-March 1934, p. 459-461.

8 Douglas Spencer "The Frankness of the Subjects on Personality Measures", in the Journal of Educational Psychology, Vol. 29, No. 1, issue of January 1938, p. 26-35.

9 Arthur W. Combs, "A Comparative Study of Motivations as Revealed in Thematic Apperception Stories and Autobiographies", in the Journal of Clinical Psychology, Vol. 3, No. 1, issue of January 1947, p. 65-75.

Ellis,¹⁰ Wesman,¹¹ Davids,¹² and, Davids and Pildner¹³ have supported Allport's postulate.

A direct or non-projective measure of personality can be operationally defined then, as one which elicits similar information about motives and attitudes from a group of normal individuals as from a group of neurotic or maladjusted persons. Likewise, a projective measure of personality is one that elicits similar information about motives and attitudes in reference to the same objects of inquiry as a direct measure with a group of normal subjects, but differs from the results of the direct measure with a group of neurotic, maladjusted individuals, in that it shows less worthy motives and poorer attitudes.

B. General Hypotheses

If the process of responding to any sentence completion item can be explained in three stages - (a) the stem acts as a stimulus to which the subject reacts inwardly, (b) the

¹⁰ Albert Ellis, "A Comparison of the Use of Direct and Indirect Phrasing in Personality Questionnaires", in Psychological Monographs, Vol. 61, No. 3, (Whole No. 284), of 1947, 41p.

¹¹ A.G. Wesman, "Faking Personality Test Scores in a Simulated Employment Situation", in the Journal of Applied Psychology, Vol. 36, No. 2, issue of April 1952, p.112-113.

¹² A. Davids, "Comparison of Three Methods of Personality Assessment, Direct, Indirect, and Projective", in the Journal of Personality, Vol. 23, No. 4, issue of June 1955, p.423-440.

¹³ Anthony Davids and Henry Pildner, "Comparison of Direct and Projective Methods of Personality Assessment Under Different Conditions of Motivation", in Psychological Monographs, Vol.72, No.11, (Whole No.464) of 1958, 30p.

latent reaction is checked against the content of the stimulus and against the situation of the subject, (c) the verbalized response or manifest reaction ^{may} represent a verbalization of the latent reaction or of the reaction altered or disguised - then the following hypotheses may be expressed.

1. In responding to the self-referent stimulus item the subject is completely aware that his response refers to himself and that another person may make an appraisal of him from his responses. Therefore, it is hypothesized that the self-referent form of a sentence completion test is a direct or non-projective instrument of personality appraisal.

2. In responding to the other-referent stimulus item the subject assumes that his response will be devoid of personal reference. Hence, he feels that he is protected from appraisal by other persons and tends to reveal latent levels of behaviour which he might not be inclined to reveal on a self-referent stimulus item for the same object of inquiry. Therefore, it is hypothesized that the other-referent form of a sentence completion test is a projective instrument of personality appraisal.

On the basis of the rationale proposed for this study, these hypotheses were reformulated in negative terms to facilitate the use of statistical methods in the analysis of the results of the experiment:

1. There is no significant difference between the self-referent form of the sentence completion items and the other-referent form for a normal, well-adjusted group.

2. There is no significant difference between the self-referent form of the sentence completion items and the other-referent form for a maladjusted group.

3. There is no significant difference between a well-adjusted group and a maladjusted group on the self-referent form of the sentence completion items.

4. There is no significant difference between a well-adjusted group and a maladjusted group on the other-referent form of the sentence completion items.

If, as has been hypothesized, the self-referent form is a direct measure and the other-referent form is a projective measure or personality assessment, then it would be expected that no significant differences would be found between the adjusted and maladjusted groups on the self-referent form, and between the self-referent and other-referent forms for the adjusted group. However, significant differences would be expected between the adjusted and maladjusted groups on the other-referent form in that the maladjusted group would obtain higher negative scores. Also the maladjusted group would obtain significantly higher other-referent scores than self-referent scores.

2.- The Sample Groups

In the selection of the two groups of "adjusted" and "maladjusted", it was essential that the criterion estimate be independent of the tests under consideration. The criterion situation had to be as homogeneous as possible with respect to external stress. It had to confront the subjects with common experiences to which they were expected to adjust. Again, the criterion had to be a psychiatric criterion, that is, a concurrent independent psychiatric examination to evaluate the subjects' current psychiatric status. Further, it was not intended nor was it desirable that the criterion be one of taxonomic diagnosis.

A. The Nature of the Criterion Used

The aim was to obtain two groups differing in socio-emotional adjustment with the variance attributable to other factors reduced to a minimum. It was the intention that these groups should be selected in such a way as to reflect extremes in adjustment to maximize the likelihood of obtaining significant differences between groups, thus giving an upper criterion, or "adjusted" group, and a lower criterion, or "maladjusted" group.

In their paper, Raines, Wittson, Hunt and Hermann¹⁴ reported:

Our findings demonstrate not only that a continuum of maladjustment exists but also that psychiatrists can recognize this continuum and make valid judgments concerning it. Basic to the practice of all neuro-psychiatric selection are three fundamental assumptions: that adjustment exists on a continuum ranging from poor to good, that trained psychiatrists are able to place a man in his position in this continuum, and that from this placement valid predictions can be made concerning the man's future psychiatric behaviour in the Naval Service. Repeated researches raise these hypotheses from the level of logical assumption to that of demonstrated fact.¹⁵

Psychiatric screening, which is concerned with the present adjustment of an individual as distinct from psychiatric prognostic selection, is based on the first and second of the above assumptions.

Hunt,¹⁶ in a later paper, pointed out that psychiatric screening is concerned with the emotional adjustment of an individual in terms of his resistance to frustrating environmental conditions, presence or absence of diagnosable mental disorders, presenting symptomatology, healthy or unhealthy personality structure, and so on. He stated a

¹⁴ George H. Raines, Cecil L. Wittson, William A. Hunt and Robert S. Hermann, "Psychiatric Selection for Military Service", in The Journal of the American Medical Association, Vol. 156, Issue of October 1954, p.817-821.

¹⁵ Ibid.

¹⁶ William A. Hunt, "A Rationale for Psychiatric Selection", in The American Psychologist, Vol. 10. No. 5, issue of May 1955, p.199-204.

further assumption to be "that certain tasks or endeavours will involve environmental pressure with which the maladjusted individual will be unable to cope, with a resulting decrement or failure in performance".¹⁷

B. The Need for Extreme Groups in Research Studies

Hunt, Herrmann and Noble¹⁸ in their study of the psychiatric interview concluded that a psychiatrist can apparently separate a "normal" group from a "borderline" group but cannot make such clear-cut distinctions within the "borderline" group. The implication drawn from this, is that in any research study where the criterion involves the use of the psychiatric interview to select "adjusted" and "maladjusted", "normal" or "neurotic" subjects, extreme groups should be used.

It may well be, as Allport contends, that an individual is either "normal" or "neurotic". A continuum of maladjusted may appear to exist because of the insensitivity of present instruments to discriminate within the so-called borderline group. For the present study, it made little difference if sides would be taken with Allport or with Hunt, et al. The main concern was to select two groups: the one "maladjusted", the other "adjusted", according to the criterion defined operationally.

¹⁷ Ibid.

¹⁸ William A. Hunt, Robert S. Herrmann and Howard Noble, "The Specificity of the Psychiatric Interview", in the Journal of Clinical Psychology, Vol.13, No.1, issue of January 1957, p.49-53.

C. The Population

The population from which the two groups were drawn consisted of 647 male, English-speaking recruits entering the Royal Canadian Air Force between January and August 1959. From this population two groups were drawn: a "normal" and a "maladjusted" group.

Fifty normal subjects were drawn at random from those of the population that on psychiatric assessment were classed as being capable of carrying out both training and trade satisfactorily and who later successfully completed the nine-week period of recruit training.

The "maladjusted" group consisted of those of the population that were released from the service for reasons of socio-emotional maladjustment. This group included twenty-nine subjects.

All of the subjects met the Air Force requirements with regard to age, marital status, citizenship, intelligence, education and physical fitness. The normal group ranged in age from seventeen to thirty years. Eighty-six percent were between seventeen and twenty-one years of age, and fourteen percent were twenty-two years and over. The maladjusted groups ranged in age from seventeen to twenty-seven years. Eighty-six point three percent were between seventeen and twenty-one years and thirteen point eight percent were twenty-two years and over. Two of the normal group were married; the remainder

were single. None of the maladjusted group were married. All of the subjects were English-speaking Canadian citizens or British subjects resident in Canada prior to their enlistment. Intelligence, as measured by the RCAF Classification Test, placed all the subjects in the upper forty percent of all applicants for the R.C.A.F.¹⁹ All the subjects were of grade eight standing or above in education.

D. The Nature of the Criterion Situation

To a civilian entering the armed services, the first problem encountered is the military situation itself. The loss of the home environment and of protective identification figures of parents, older siblings, teachers, employers and the like, present a need for new identifications. Additional stress are represented by the military discipline, the loss of personal privacies, and the regimented routine. The stress demand adjustment. For some, the airman or airwoman recruit status represents a change of social status and may threaten the individual's feelings of security.

¹⁹ A.C. Keating and K. Jonassohn, Classification Test Scores of Applicants for the RCAF: 1 December 1957 to 28 February 1959, RCAF Personnel Standards Analysis Branch Technical Report No. 59-3, Ottawa, 1959, 10p. In this report, it was noted that nearly sixty percent of the 21,773 applicants for the RCAF groundcrew in the period of time covered by the report were below the acceptable standard.

The person who is unable to adjust to the military situation, to authority, or to the newly acquired status is typically either immature in some important aspect of development or psychopathic. Immaturity is seen frequently in unrealistic superficial motivation, lack of definite goals, narcissistic need for self-gratification, with consequent lack of frustration tolerance, neurotic sexual fixation at a pre-heterosexual level, and a juvenile emotionality. Psychopathic behaviour leads the individual to infractions of regulations, irresponsible and careless behaviour, and sometimes serious friction with authority figures.

In general, the military training situation was thought to be sufficiently stressful to identify many of the immature, neurotic, and psychopathic components of the training population.

3.- Procedure

The stimulus items of the Sack's Sentence Completion Test were reworded by placing third personal pronouns in place of first personal pronouns. The original items and the reworded items were then pooled and drawn at random to form a sentence completion test of 120 stimulus items. Sixty of these were self-referent and sixty were matched other-referent items.

A. Test Administration

The Experimental Sentence Completion Test was administered routinely as part of the Airmen Classification Battery of Tests at the Personnel Selection Unit at the Royal Canadian Air Force Station St. Johns, Quebec. Subjects were informed that the sentence completions should be completed within thirty minutes, but that extra time would be allowed if required. All English-speaking male recruits processed through the Selection Unit between January and August 1959 received the test. Following each test administration the test forms were sent directly to the experimenter for storage. By that means the psychiatric assessments were uncontaminated by knowledge of the Experimental Sentence Completion Test.

Following the period of test administration the subjects for the groups were selected according to the criteria given above. Their test forms were then withdrawn from those of the population and mixed haphazardly by a person other than the experimenter.

B. Scoring Procedures

In order to provide for the objective scoring of the responses to the Experimental Sentence Completion Test a scoring manual²⁰ was written. For each item in the test there

²⁰ In format this manual was based on an unpublished Provisional Scoring Manual for the S.A.M. Sentence Completion Test loaned to the experimenter by Dr. S.B. Sells, formerly of the United States Air Force School of Aviation Medicine.

was assembled a list of completions related to the applicable scoring categories.²¹ Self-referent and the matching other-referent items were scored from the same examples. The process of scoring consisted in rating each completion either "positive" or "negative" using the instructions and the examples given in the scoring manual as a guide. Blank items were not scored. The basic score unit was the "negative" response. To score each protocol an Other-Referent Rating Record Sheet and a Self-Referent Rating Record Sheet were used. On those sheets, the item numbers were grouped according to the four attitude areas covered by the test: Family, Sex, Interpersonal Relations, and Self-Concept. The items rated in the order given by the Rating Sheets. With each protocol the Other-Referent items were scored first.

C. Interscorer Reliability

In order to achieve maximum consistency all the protocols were scored by the same person. This presented the problem of investigating the extent to which the scoring would have been influenced by the subjective judgment of the scorer. To obtain two measures of interscorer reliability,

²¹ To obtain characteristic "positive" and "negative" completions the experimenter obtained the manuals for seven widely-used sentence completion tests. He then grouped the items according to the similarity of their content and used the examples of "positive" and "negative" completions given by their authors. Only items that were quite similar to those in the Sack's Sentence Completion Test were used.

twelve protocols from subjects not included in the study were scored independently by the scorer and another psychologist. Coefficients of correlation between the scores assigned by the scorer and the other psychologist for the Other-Referent and Self-Referent Forms of the test were calculated. These values are presented in the next chapter.

D. Test Reliability

To provide measures of the degree of consistency of the test, product-moment coefficients of correlation between two halves²² of each form of the test were computed, using thirty protocols from the adjusted group. Both coefficients of correlation were corrected by the Spearman-Brown prophesy formula to give coefficients of reliability for the Self-Referent Form and the Other-Referent Form of the test. Now the Spearman-Brown prophesy formula is based on the assumption that all of the items in the test measure the same factor. If more than one factor is involved, and this was highly probable with the type of test being used, that method gives an underestimate of the reliability coefficient. Thus, the experimenter felt justified in using that procedure, in that it was better to underestimate the reliability coefficients than to overestimate them.

22 The halves were obtained by taking the first two items on the Rating Record Sheet for the first half, the second two items for the second half, the third pair of items for the first half, and the fourth pair of items for the second half,

4.- Statistical Treatment of the Data

The t Test and the F Test were employed to test the significance of the differences between the means and the variances of the other-referent and self-referent scores for each category of items within each group. The assumption was made that the trait measured by each form of the test is normally distributed. Also, it was assumed that the two sets of data within each group were correlated, so that the statistical formulae appropriate to this assumption were used.

The t Test was employed to test the significance of the differences between the means of the adjusted and maladjusted groups on each form of the Experimental Sentence Completion Test. Again the assumption was made that the trait measured by each form of the test is distributed normally, but the formula for independent means was used.

In the following chapter the results of the study will be presented, analyzed and discussed. The findings within both the adjusted and maladjusted groups will be presented and comparisons made between the two groups.

CHAPTER V

FINDINGS AND DISCUSSION

Both the reliability of the scoring procedure and of each form of the test, together with the statistics descriptive of the behaviour of the two groups on the test, will be reported in this chapter. From the analysis of these statistics,¹ inferences related to the four null hypotheses, set out in the preceding chapter, will be drawn and discussed.

1.- Interscorer Reliability

Twelve protocols taken from outside of the two groups were scored independently by the experimenter and by another with clinical experience. The product-moment coefficient of correlation between the two sets of scores for the Other-Referent Form of the test was .934. The interscorer reliability index for the Self-Referent Form was .966. Both of these coefficients of correlation are significant beyond the 1 per cent point. The scoring procedures were considered to be sufficiently reliable to permit statistical analyses of the results obtained by them.

¹ In the statistical analysis of the data a level of significance of 1 per cent was the criterion chosen to reject the possibility that a result might have occurred by chance. A level of significance between the 5 and the 1 per cent was taken as being suggestive but as requiring further experimentation.

2.- Test Reliability

Using thirty pairs of test rating sheets obtained from the "adjusted" group, reliability coefficients were calculated in the manner described in Chapter IV. The corrected reliability coefficient for the Other-Referent Form was .875; that for the Self-Referent Form was .900. Both are significant beyond the 1 per cent level. They were considered to be underestimates of the true reliability coefficients as the correction formula assumes that all the items in a given test measure the same factor. They appeared to be at the upper end of the range of reliability indices presented by various authors for other sentence completion tests. Hence, it was concluded that both forms of the test could be considered as relatively stable.

3.- Null Hypotheses One and Two

The first hypothesis was stated: There is no significant difference between the self-referent form of the sentence completion items and the other-referent form for a normal, well-adjusted group. Data pertaining to the adjusted group are presented in Table I. It will be seen that the values of t for the total and for each category within the tests reached a level of significance great enough to render the null hypothesis highly improbable. It was thought that this discrepancy might have been due to a difference in variability rather than to a difference in the means.

Table II presents a comparison of the variances within the normal group on each of the test forms. The variances for the categories "Sex" and "Interpersonal" differ significantly, whilst the F value for the "Total" approaches the value of F required for the five per cent level of significance. For the categories "Sex" and "Interpersonal", the scores on the self-referent items obtained by the adjusted group tend to cluster around their means to a greater extent than do the scores obtained in the other-referent items.

Applying the null hypothesis to each category of items in turn we find that for the categories "Interpersonal" and "Sex" there are significant differences in the variances on the two types of items, namely, self-referent and other-referent. We can make no positive statement with regard to the significance of the difference between the means for each of these categories, since the t test assumes variances to be equal.

In the case of the categories "Family" and "Self", the null hypothesis can be rejected, and it can be stated that there are significant differences between the self-referent form of the sentence completion items and the other-referent form for a normal, adjusted group when these items refer to the categories "Family" and "Self".

In the case of the complete test, the variance differences were not shown to be statistically significant,

Table I.- Means and Differences between Means of Negative Other-Referent and Self-Referent Scores for the Adjusted Group.

Category of Items	Other-Referent Mean	Self-Referent Mean	Difference between Means ^a	t
Interpersonal	4.34	2.16	2.18	5.806
Sex	3.24	1.64	1.60	6.301
Family	4.72	2.44	2.28	6.603
Self	8.96	7.22	1.74	4.108
Total	21.26	13.46	7.80	8.657

^a All the differences were significant at the one per cent level of probability.

Table II.- Comparison of the Variances within the Adjusted Groups on the Self-Referent and Other-Referent Items.

Category of Items	Variance (S^2)		F
	Other-Referent Items	Self-Referent Items	
Interpersonal	9.0644	3.9744	2.28 ^a
Sex	4.1424	2.1104	1.96 ^a
Family	6.7616	4.4464	1.52
Self	11.0384	9.8116	1.13
Total	77.9124	49.2084	1.58 ^b

^a The variabilities differ significantly at the one per cent level of probability.

^b This value approaches the value of F required for the five per cent level of significance.

although further research might be warranted since F approached the five per cent level of significance. Hence, the means may be taken to be statistically significant. The null hypothesis can be rejected, and it can be stated that there are significant differences between the Self-Referent Form of the Experimental Sentence Completion Test and the Other-Referent Form of the same test for an adjusted group.

The second hypothesis was stated: There is no significant difference between the self-referent form of the sentence completion items and the other-referent form for a maladjusted group. Data pertaining to the maladjusted group is presented in Table III. It will be seen that the value of t for the total and for each category but one ("Sex") within the tests, reached a level of significance great enough to render the null hypothesis highly improbable. The F test showed that the variances differed significantly in only one category of items, namely, "Sex". In this category scores tended to cluster to a greater extent around the mean of the self-referent form of the items than around the mean of the other-referent type.

Justification is felt for rejection of the second hypothesis when applied to the categories "Interpersonal", "Family", "Self", and "Total". It may be stated that significant differences between self-referent and other-referent forms of items having reference to interpersonal relations,

Table III.- Means and Differences between Means of Negative Other-Referent and Self-Referent Scores for the Maladjusted Group.

Category of Items	Other-Referent Mean	Self-Referent Mean	Difference between Means ^a	t
Interpersonal	5.07	3.03	2.04	3.751
Sex	3.14	2.03	1.11 ^b	2.613
Family	4.76	2.45	2.31	6.107
Self	10.31	8.66	1.65	2.889
Total	23.28	16.17	7.11	5.285

^a All differences were significant at the one per cent level of probability except where noted as being otherwise.

^b Significant at the five per cent level of probability.

Table IV.- Comparison of the Variances within the Maladjusted Group on the Self-Referent and Other-Referent Items.

Category of Items	Variance (S^2)		F
	Other-Referent Items	Self-Referent Items	
Interpersonal	12.133	10.653	1.14
Sex	6.325	2.240	2.82 ^a
Family	9.769	6.730	1.45
Self	12.972	14.639	1.13
Total	121.717	93.315	1.30

^a The variabilities differ significantly at the one per cent level of probability.

the family, and the self were found within the maladjusted group. Generally speaking, it may be stated that there are significant differences between the Self-Referent Form of the Experimental Sentence Completion Test and the Other-Referent Form of the same test for a maladjusted group. Further study of items related to sex adjustment is warranted before the specific conclusions related to them could be made.

4.- Null Hypotheses Three and Four

In the previous chapter, the third null hypothesis was stated: There is no significant difference between a well-adjusted group and a maladjusted group on the self-referent form of the sentence completion items. Data showing the comparisons of the two groups on the Self-Referent Form of the Experimental Sentence Completion Test are presented in Table V. Since none of the differences between the means of the adjusted and maladjusted groups was found to be statistically significant, no justification for the rejection of the null hypothesis was found.

The fourth hypothesis was stated: There is no significant difference between a well-adjusted group and a maladjusted group on the other-referent form of the sentence completion items. Data showing the comparisons of the two groups on the Other-Referent Form of the Experimental

Sentence Completion Test are presented in Table VI. Since none of the differences between the means of the adjusted and maladjusted groups was found to be statistically significant, no justification for the rejection of this fourth null hypothesis was found.

5.- General Discussion

Referring back to the problem outlined in the first part of the previous chapter, three specific question were proffered:

1. Do responses of individuals to self-referent and other-referent stimuli in a sentence completion test differ significantly in terms of the amount of unfavourable attitudes, disturbances, and problems revealed?

2. Is the self-referent type of stimulus item more revealing of maladjustment than the other-referent form of the same item or vice versa?

3. Are both types of items "projective" techniques?

Answers to these questions, based on the study made, can now be advanced. Generally, it was found that responses of individuals, both adjusted and maladjusted, to self-referent and other-referent stimuli in sentence completion tests do differ significantly in that other-referent items tend to bring forth the expression of more unfavourable attitudes, disturbances, and problems. Sacks had noted

Table V.- Means and Differences between Means of the Maladjusted and Adjusted Groups on the Self-Referent Form of the Experimental Sentence Completion Test.

Category of Items	Maladjusted Group Mean	Adjusted Group Mean	Difference between Means ^a	t
Interpersonal	3.03	2.16	0.87	1.287
Sex	2.03	1.64	0.39	1.123
Family	2.45	2.44	0.01	0.014
Self	8.66	7.22	1.44	1.688
Total	16.17	13.46	2.71	1.302

^a None of the differences between the means are statistically significant.

Table VI.- Means and Differences between Means of the Maladjusted and Adjusted Groups on the Other-Referent Form of the Experimental Sentence Completion Test.

Category of Items	Maladjusted Group Mean	Adjusted Group Mean	Difference between Means ^a	t
Interpersonal	5.07	4.34	0.73	0.927
Sex	3.14	3.24	0.10	0.108
Family	4.76	4.72	0.04	0.054
Self	10.39	8.96	1.43	1.723
Total	23.28	21.26	2.02	0.827

^a None of the differences between the means are statistically significant.

this to be so particularly in attitudes expressed toward the father and toward superiors.¹ This might explain why Symonds, reporting on the O.S.S. study,² could find no differences between a highly adjusted and a low adjusted group.

Some underlying hypothetical constructs are required to explain this difference between type of items. One construct might be that third-personal pronouns are literally referred to other persons actually living. If such were the case, why should the subjects pick on others who were more maladjusted than themselves?

It has been pointed out elsewhere that "Self theorists have held, in one form or another, that the basic human need is the preservation and enhancement of the self".³ Attributing to an unknown third person expressions of maladjustment could well be a means of satisfying this need.

With reference to the second question, it cannot be stated that one form of the items is more revealing of actual maladjustment than the other, since the "adjusted" and "maladjusted" groups did not differ significantly on either form

1 J.M. Sacks, op. cit.

2.P.M. Symonds, op.cit., p.238.

3.Philip Worchel, Adaptability Screening of Flying Personnel: Development of a Self-Concept Inventory for Predicting Maladjustment, Report No. 56-62, Randolph Field (Texas), U.S.A.F. School of Aviation Medicine, 1957, p.2.

of the items. Calvin and Holtzman⁴ found on the basis of a sociometric scale, and a self-rating, that the tendency to enhance the self is inversely related to maladjustment. On the basis of their findings it could have been expected that the self-referent form of the item would reveal differences between the groups.

Using Allport's postulate that direct and projective measures of assessment will reveal the same information about the motives and attitudes of the normal, well-adjusted individual, but that with the maladjusted person projective techniques will elicit repressed material which is contrary to that elicited on direct measures of personality - this study has not shown either form of the sentence completion test item to be a projective technique.

Some comments upon the experiment reported seem necessary. First, the experimenter in selecting the control and experimental groups avoided the ex post facto psychiatric criterion, so often found in other studies, by having the subjects tested prior to their psychiatric assessment. Secondly, contamination of the psychiatric criterion by the test results was rigourously avoided by having the tests packaged immediately after their having been given, and sent

⁴ A.D. Calvin and W.H. Holtzman, "Adjustment and the Discrepancy between Self-Concept and Inferred Self", in the Journal of Consulting Psychology, Vol. 17, issue of 1953, p. 39 - 44.

to the experimenter for storage. Scoring contamination by the rater having knowledge of the group to which the subject belonged was avoided as far as possible by having another person mix the tests of the two groups together in random fashion.

Some criticism of the scoring techniques used could be offered. In the first place, the scorer had to rely heavily on samples of completions suggested by other authors, in order to assess the responses being rated. Again, it was not possible to allow for differences in "maladjusted" responses - all had a weight of one, whereas, in actuality some responses were more suggestive of maladjustment than others.

In conclusion, it should be emphasized that the subjects were tested in a "real-life" situation, and were not placed in an artificial situation of the experimenter's own construction.

SUMMARY AND CONCLUSIONS

An experiment was carried out to compare the self-referent sentence completion item with the other-referent type, and to assess objectively, whether these types of items could be classed as "projective" techniques. The rationale for the study was Allport's postulate that direct and projective measures of assessment would reveal the same information about the motives and attitudes of the well-adjusted person, but with the maladjusted individual, projective techniques would elicit repressed material which would be contrary to that expressed on direct measures of personality.

Two experimental groups differing in socio-emotional adjustment as assessed by a psychiatric interview, in similar environments, were selected from a population of male, air force recruits. These had already been tested by an experimental form of the Sacks' Sentence Completion Test. The latter was made up of one hundred and twenty sentence beginnings. Sixty of the items contained a first personal pronoun; the remainder were matched items containing a third personal pronoun. These items were mixed in random order. Each set of sixty items constituted a "test". Completions were scored as being "positive" or "negative" with respect to the adjustment indicated. Scores for each "test" consisted of the total number of responses classed as being "negative".

It was hypothesized that:

1. There would be no significant differences between the two forms of the Sentence Completion Test for either group.

2. There would be no significant differences between groups on either form of the Sentence Completion Test.

On the basis of the results obtained, the first null hypothesis was rejected in the case of each group. The second hypothesis could not be rejected in the case of either form of the test. It was suggested that whilst other-referent items tended to elicit a greater number of "negative" responses, neither type of item could be considered to be "projective" under the rationale stated for the study. Also, it was impossible to separate the adjusted group from the maladjusted group on the basis of their test scores with better than chance success. The present results, with respect to the validity of the instrument, suggest that the utmost care should be taken to validate such tests in the situation in which they are to be used, rather than to rely on validities reported elsewhere.

Further research is required utilizing items referring to the category "sex". Likewise, items from other sentence completion tests might be utilized. Matched self-referent and other-referent items in multiple-choice form might be compared. Different criteria of adjustment-maladjustment could be utilized in such studies.

Research into personality theory is required. It is not sufficient to note that other-referent stems elicit more unfavourable attitudes, disturbances, and problems. Why is this so? What is the theoretical significance of such behaviour?

BIBLIOGRAPHY

Allport, G.W., "The Trend in Motivational Theory", in the American Journal of Orthopsychiatry, Vol. 23, No. 1, issue of 1953, p. 107-119.

Allport discusses the unreasonable separation of direct and projective measures of personality, the tendency to ignore direct measures in favour of the projective techniques, and the trend of motivational theory toward irrationality and geneticism, and away from accepting healthy, adult, psychogenic motives at their face value. He further stresses the need for the complementary usage of both direct and projective techniques in order that projective measures can receive proper interpretation and that a maximally accurate personality picture be gained. He postulates that direct and projective measures will reveal the same information about the motives of the normal, well-adjusted personality. With neurotics, however, projective techniques will elicit repressed material and information which is contrary to that obtained from direct measures. Recent experimental literature has supported this postulate.

Cromwell, Rue L. and Richard M. Lundy, "Productivity of Clinical Hypotheses on a Sentence Completion Test", in the Journal of Consulting Psychology, Vol. 18, No. 6, issue of December 1954, p. 421-424.

A sixty-five item sentence completion test was administered to sixty neuropsychiatric hospital admissions. The order of the stems and two different sets of instructions were alternated from subject to subject. Thirty-nine clinical psychologists made personality inferences from the protocols, and listed the completed sentences used in making inferences. One of the findings reported suggested that many of the more productive stems were of the self-referent type, whilst many of the poorer stems were of the other-referent or impersonal type. This study adds support to Sacks.

Davids, Anthony, "Comparison of Three Methods of Personality Assessment, Direct, Indirect, and Projective", in the Journal of Personality, Vol. 23, No. 4, issue of June 1955, p. 423-440.

The author reports on a study designed to compare measures of effectiveness of adjustment derived from the hypothesized three methods of personality assessment: direct, indirect and projective. The study substantiated Allport's assertion that the personality test performance by the normal

person will be consistent in direct and projective measures. The study was well designed and reported. Some weaknesses of the study were corrected by a later investigation.

----- and Henry Pildner, "Comparison of Direct and Projective Methods of Personality Assessment under Different Conditions of Motivation", in Psychological Monographs, General and Applied, Vol. 72, No. 11, (Whole No. 464), issue of 1958, 30 p.

These authors report on a more extensive study of the relative performance on direct and projective personality assessment measures by subjects under different motivational conditions. Two groups of subjects were studied. Five direct and four projective measures were utilized. The results demonstrated that distortion is likely to take place on direct personality measures as the result of the motivational situation. Likewise, projective methods can be used effectively to control assessment deception. In the sentence completion test used by the authors, fifty items were of the self-referent type and fifty were of the other-referent type. The authors concluded that self-referent items were not as efficient as other-referent stems in eliciting unconscious or inhibited material, and the responses to them could be distorted. This is an example of a well-planned study more in the area of social psychology than in clinical psychology.

Dorris, Ronald J., Daniel J. Levinson and Eugenia Hanfmann, "Authoritarian Personality Studied by a New Variation of the Sentence Completion Technique", in the Journal of Abnormal and Social Psychology, Vol. 49, No. 1, issue of January 1954, p. 99-108.

As part of their study of the Authoritarian Personality, the authors found that other-referent items elicited significantly more denials of self-reference than did self-referent items. They concluded that this finding was consistent with the frequently made assumption that other-referent items enable the projection of tendencies that the subject does not recognize in himself.

Forer, Bertram R. and Ruth S. Tolman, "Some Characteristics of Clinical Judgment", in the Journal of Consulting Psychology, Vol. 16, No. 5, issue of October 1952, p. 347-352.

A group of clinical psychologists rated the 100-item Forer Sentence Completion Test Blank in terms of the degree of potential clinical value. One finding was that no significant differences in value judgments existed between self-referent and other-referent items within the same content area. From the study, it is not clear as to whether

fifty per cent of the clinical psychologists preferred the one type and the remainder preferred the other, or if the group of psychologists as a whole found both types of items useful.

Getzels, J.W., The Assessment of Personality and Prejudice by the Method of Paired Direct Projective Questionnaires, unpublished Doctoral thesis presented to the Department of Social Relations of Harvard University, Cambridge, 1951.

A well-designed study utilizing two experiments, one dealing with personality differences between twenty-five well-adjusted and forty maladjusted veterans who were applying for financial grants, and the other a study of prejudice among female college students. From the results, the author concluded that self-referent items on a sentence-completion test are not truly projective stimuli. Unfortunately, in the first experiment the author had not one experimental variable but two. Again, he used highly structured sentence stems, which are not found in most clinical tests. Thus, his results cannot be applied too rigidly to tests used in a clinical situation.

----- and J.J. Walsh, "The Method of Paired Direct and Projective Questionnaires in the Study of Attitude Structure and Socialization", in Psychological Monographs, General and Applied, Vol. 72, No. 1, (Whole No. 454), issue of 1958, 34 p.

These authors used paired sentence completion stems, that were identical except that the first personal pronoun was used in one and a proper name in the other, to study attitude structure and socialization. They assumed the former to be a direct instrument, and the latter to be a projective instrument. The study was well organized but criticism might be voiced of their use of the term "questionnaire", and their use of items with proper names rather than third personal pronouns.

Hanfmann, Eugenia and J.W. Getzels, "Studies of the Sentence Completion Test", in the Journal of Projective Techniques, Vol. 17, No. 3, issue of September 1953, p. 280-294.

A study of the sources of reference in responses to sentence completion items revealed that a large percentage of responses to other-referent items was determined by unconscious self-reference.

Kimball, Barbara, "The Sentence Completion Technique in a Study of Scholastic Underachievement", in the Journal of Consulting Psychology, Vol. 16, No. 5, issue of October 1952, p. 353-358.

Kimball gave fifty other-referent incomplete sentences to a group of underachievers and a control group. The items related to "father", and to "aggression toward living things", distinguished between the two groups. This study gives little insight into the underlying rationale of the items used, but adds to the understanding of the personality dynamics of the scholastic underachiever.

Meltzoff, Julian, "The Effect of Mental Set and Item Structure upon Response to a Projective Test", in the Journal of Abnormal and Social Psychology, Vol. 46, No. 2, issue of April 1951, p. 177-189.

The author describes an experiment to determine the effects of the subject's mental set, and the tone or attitude expressed in the stimulus items, upon responses to a sentence-completion test. This study contributes to the understanding of how a subject behaves on a sentence completion test when different instructions or different items are used.

Meyer, Mortimer M. and Ruth S. Tolman, "Parental Figures in Sentence Completion Test, in TAT and in Therapeutic Interviews", in the Journal of Consulting Psychology, Vol. 19, No. 3, issue of June 1955, p. 170.

The authors compared the completions of twenty patients to items referring to "mother" and "father" in the Forer Sentence Completion Test with the attitudes expressed in TAT and therapeutic sessions. It was concluded that the attitudes expressed in sentence completions furnish no basis of prediction of those found in either TAT stories or in therapy sessions. This study was not well designed, and was based on too small a sample to furnish conclusive evidence. Its value is in its suggestion for further research.

Rohde, Amanda R., The Sentence Completion Method, Its Diagnostic and Clinical Application to Mental Disorders, New York, Ronald Press, 1957, xii-301p.

This manual provides the most complete published review of the literature dealing with the sentence completion method, together with a lengthy bibliography. The weakness of the review is that it does not attempt to be critical. Some attention is given to the underlying rationale of these tests. Three chapters are given to describing the administration, scoring, standardization, and interpretation of the Rohde Sentence Completion Method. Four chapters deal with differential diagnosis and clinical application. A valuable part of the manual contains twenty-three protocols scored and interpreted.

Sacks, J.M., "The Relative Effect upon Projective Responses of Stimuli Referring to the Subject and of Stimuli Referring to Other Persons", in the Journal of Consulting Psychology, Vol. 13, No. 1, issue of February 1949, p. 12-20.

A report of research conducted with a view to investigating the validity of the assumption that a subject revealed more about himself when he talked about other people or about impersonal unstructured objects than when he talked about himself. In the experiment, Sacks compared self-referent with other-referent stimuli; the latter containing proper names. Six of the seven psychologists rating the protocols found the self-referent items to be superior. Unfortunately, the author used a heterogenous group of neurotic and psychotic patients, and no normal group. The validity of the use of the Chi Square technique in the analysis of the results, is dubious. However, the study is a landmark in the history of the sentence completion test, in that it draws attention to the lack of experimental evidence to support various underlying hypotheses.

----- and Sidney Levy, "The Sentence Completion Test", in L.E. Abt and L. Bellak, Ed., Projective Psychology, Clinical Approaches to the Total Personality, New York, Knopf, 1950, p. 357-402.

Essentially the manual for the Sacks Sentence Completion Test, it provides examples for scoring the completions, and samples of the record form.

Stein, M.J., "Use of a Sentence Completion Test for the Diagnosis of Personality", in the Journal of Clinical Psychology, Vol. 3, No. 1, issue of January 1947, p. 47-56.

This article describes the test used in the Office of Strategic Service selection programme. Fifty items referred to the behaviour, attitudes and activities of third parties; fifty others were of the self-referent type. All were mixed in random order. The author considered the other-referent items to be the more projective, whilst the test as a whole offered the clinician clues as to critical areas in the personality, and some knowledge of the dynamics involved. This is a subjective appraisal and no objective data are presented to give weight to the author's opinions.

Symonds, P.M., "The Sentence Completion Test as a Projective Technique", in the Journal of Abnormal and Social Psychology, Vol. 42, No. 3, issue of July 1947, p. 320-329.

This report describes an attempt to obtain validation data with the Office of Strategic Service test. Considering the somewhat inadequate criterion used, the author is overly

critical of the test in his conclusions. The worth of the report is that it draws attention to the much needed validation of the many clinical tests of which the sentence completion test is one.

Trites, David K., Psychiatric Screening of Flying Personnel, Evaluation of Assumptions underlying Interpretation of Sentence Completion Tests, Report Number 55-33, Randolph Field (Texas), U.S.A.F. School of Aviation Medicine, 1955, 9p.

A factorial study of an 88-item test was made by Trites to evaluate two assumptions often held in the objective interpretation of sentence completion tests. Both of the assumptions were partially supported by the data. Four factors were produced: the Self-Centred Anxiety Factor, the Air Force Motivation Factor, the Interpersonal Factor, and the Narcissistic (or Self-Enhancement) Factor. The two latter factors seemed best defined by self-referent items. The results are an original contribution to the knowledge of this area.

-----, "Evaluation of Assumptions underlying Interpretation of Sentence Completion Tests", in the Journal of Consulting Psychology, Vol. 20, No. 1, issue of February 1956, p.8.

This is a summary of the author's earlier report.

APPENDIX 1

EXPERIMENTAL

SENTENCE COMPLETIONS, FORM M

THIS MATERIAL IS CONFIDENTIAL AND WILL BE USED FOR RESEARCH ON ATTITUDES TOWARDS "LIFE IN GENERAL". IT WILL NOT BE READ HERE NOR WILL IT BE PLACED ON YOUR CONFIDENTIAL FILE.

INSTRUCTIONS:

Below are one hundred and twenty partly completed sentences. Read each one and finish it by writing clearly the first thing that comes into your mind. Sometimes you will find a single word will complete the sentence and sometimes you will find that a brief phrase will do it. A brief phrase is preferred, but if one word is all you can think of, then that will be enough. Work as quickly as you can. If you cannot complete an item circle the number and return to it later.

1. He feels that his father seldom _____
2. My idea of a perfect woman _____
3. The worst thing he ever did _____
4. I would do anything to forget the time I _____
5. When the odds are against me _____
6. Most of my friends don't know that I am afraid of _____
7. When he was a child _____
8. Before he enlisted, he _____
9. If I were young again _____

10. Some day I _____
11. His mother _____
12. If his father would only _____
13. When he was a child, his family _____
14. If I had sex relations _____
15. The people I like best _____
16. His greatest mistake was _____
17. My mother _____
18. When I see a man and a woman together _____
19. Most of his friends don't know that he is afraid of _____
20. The people who work for me _____
21. When luck turns against me _____
22. People who work with him usually _____
23. When I was younger, I felt guilty about _____
24. He thinks most girls _____
25. I always wanted to _____
26. At work, he gets along best with _____
27. My sex life _____
28. People whom I consider my superiors _____
29. I could be perfectly happy if _____
30. His fears sometimes force him to _____
31. He doesn't like people who _____
32. I feel that a real friend _____
33. His feeling about married life is _____
34. He thinks that most mothers _____
35. I like my mother but _____

36. Compared with most families, mine _____
37. When he's not around, his friends _____
38. I believe most women _____
39. In school, his teachers _____
40. He feels that his father is _____
41. He looks forward to _____
42. I like working with people who _____
43. His secret ambition in life _____
44. What he likes least about women _____
45. If people work for him _____
46. What he wants most out of life _____
47. I don't like people who _____
48. My family treats me like _____
49. Those he works with are _____
50. I wish I could lose the fear of _____
51. His most vivid childhood memory _____
52. When I am older _____
53. Most families I know _____
54. He believes that he has the ability to _____
55. In giving orders to others he _____
56. If I were in charge _____
57. When I see the boss coming _____
58. My greatest weakness is _____
59. His mother and he _____
60. To him the future looks _____
61. What I want most out of life _____

62. When he sees the boss coming _____
63. He would do anything to forget the time he _____
64. When he sees a man and a woman together _____
65. If he were in charge _____
66. When he was younger, he felt guilty about _____
67. When he is older _____
68. The worst thing I ever did _____
69. In giving orders to others I _____
70. The people who work for him _____
71. Some day he _____
72. He knows it is silly but he is afraid of _____
73. He believes most women _____
74. In school, my teachers _____
75. I look forward to _____
76. His family treats him like _____
77. He wishes he could lose the fear of _____
78. He wishes his father _____
79. When I was a child _____
80. People who work with me usually _____
81. When I was a child, my family _____
82. Before I enlisted, I _____
83. If people work for me _____
84. I feel that my father seldom _____
85. The people he likes best _____
86. Those I work with are _____
87. When luck turns against him _____

88. My secret ambition in life _____
89. He could be perfectly happy if _____
90. His idea of a perfect woman _____
91. He likes his mother but _____
92. If he were young again _____
93. My fears sometimes force me to _____
94. People whom he considers his superiors _____
95. I feel that my father is _____
96. My greatest mistake was _____
97. Compared with most families, his _____
98. When I'm not around, my friends _____
99. If my father would only _____
100. The men over me _____
101. He feels that a real friend _____
102. He always wanted to _____
103. I think that most mothers _____
104. My mother and I _____
105. At work, I get along best with _____
106. My most vivid childhood memory _____
107. The men over him _____
108. I believe that I have the ability to _____
109. Most families he knows _____
110. I wish my father _____
111. I know it is silly but I am afraid of _____
112. When the odds are against him _____

113. If he had sex relations _____
114. I think most girls _____
115. His sex life _____
116. His greatest weakness is _____
117. He likes working with people who _____
118. To me the future looks _____
119. My feeling about married life is _____
120. What I like least about women _____

APPENDIX 2

RATING RECORD SHEET

SENTENCE COMPLETIONS - OTHER-REFERENT FORM

NAME:

A. FAMILY ATTITUDES

I TOWARD MOTHER

11	P	N
34	P	N
59	P	N
91	P	N

II TOWARD FATHER

1	P	N
12	P	N
40	P	N
78	P	N

III TOWARD FAMILY UNIT

13	P	N
76	P	N
97	P	N
109	P	N

B. SEX ATTITUDES

IV TOWARD WOMEN

24	P	N
44	P	N
73	P	N
90	P	N

**V TOWARD HETEROSEXUAL
RELATIONSHIPS**

33	P	N
64	P	N
113	P	N
115	P	N

C. INTERPERSONAL ATTITUDES

**VI TOWARD FRIENDS &
ACQUAINTANCES**

31	P	N
37	P	N
85	P	N
101	P	N

VII TOWARD SUPERIORS

39	P	N
62	P	N
94	P	N
107	P	N

**VIII TOWARD PEOPLE
SUPERVISED**

45	P	N
55	P	N
65	P	N
70	P	N

IX TOWARD COLLEAGUES

22	P	N
26	P	N
49	P	N
117	P	N

<u>D. SELF ATTITUDES</u>				<u>OVER-ALL TOTALS FOR "N"</u>	
<u>X FEARS</u>				<u>RESPONSES</u>	
	19	P	N	A. FAMILY	
	30	P	N	B. SEX	
	72	P	N	C. INTERPERSONAL	
	77	P	N	D. SELF	_____
<u>XI GUILT FEELINGS</u>				TOTAL "N"	_____
	3	P	N		
	16	P	N		
	63	P	N		
	66	P	N		
<u>XII TOWARD OWN ABILITIES</u>					
	54	P	N		
	87	P	N		
	112	P	N		
	116	P	N		
<u>XIII TOWARD PAST</u>					
	7	P	N		
	8	P	N		
	51	P	N		
	92	P	N		
<u>XIV TOWARD FUTURE</u>					
	41	P	N		
	60	P	N		
	67	P	N		
	71	P	N		
<u>XV GOALS</u>					
	43	P	N		
	46	P	N		
	89	P	N		
	102	P	N		

APPENDIX 3

RATING RECORD SHEET

SENTENCE COMPLETIONS - SELF-REFERENT FORM

NAME:

A. FAMILY ATTITUDES

I TOWARD MOTHER

17	P	N
35	P	N
103	P	N
104	P	N

II TOWARD FATHER

84	P	N
95	P	N
99	P	N
110	P	N

III TOWARD FAMILY UNIT

36	P	N
48	P	N
53	P	N
81	P	N

B. SEX ATTITUDES

IV TOWARD WOMEN

2	P	N
38	P	N
114	P	N
120	P	N

V TOWARD HETEROSEXUAL RELATIONSHIPS

14	P	N
18	P	N
27	P	N
119	P	N

C. INTERPERSONAL ATTITUDES

VI TOWARD FRIENDS & ACQUAINTANCES

15	P	N
32	P	N
47	P	N
98	P	N

VII TOWARD SUPERIORS

28	P	N
57	P	N
74	P	N
100	P	N

VIII TOWARD PEOPLE SUPERVISED

20	P	N
56	P	N
69	P	N
83	P	N

IX TOWARD COLLEAGUES

42	P	N
80	P	N
86	P	N
105	P	N

<u>D. SELF ATTITUDES</u>				<u>OVER-ALL TOTALS FOR "N"</u>	
<u>X FEARS</u>				<u>RESPONSES</u>	
	6	P	N	A. FAMILY	
	50	P	N	B. SEX	
	93	P	N	C. INTERPERSONAL	
	111	P	N	D. SELF	
<u>XI GUILT FEELINGS</u>				TOTAL "N"	
	4	P	N		
	23	P	N		
	68	P	N		
	96	P	N		
<u>XII TOWARD OWN ABILITIES</u>					
	5	P	N		
	21	P	N		
	58	P	N		
	108	P	N		
<u>XIII TOWARD PAST</u>					
	9	P	N		
	79	P	N		
	82	P	N		
	106	P	N		
<u>XIV TOWARD FUTURE</u>					
	10	P	N		
	52	P	N		
	75	P	N		
	118	P	N		
<u>XV GOALS</u>					
	25	P	N		
	29	P	N		
	61	P	N		
	88	P	N		

APPENDIX 4

ABSTRACT OF

Self-Referent versus Other-Referent Incomplete Sentences in Projection¹

1.- Statement of the Problem

In the past, investigators have failed to reach agreement as to the projective or non-projective nature of those sentence completion stimulus items worded to include a first personal pronoun, and those worded to include a third personal pronoun.

It was the purpose of this study to seek answers to three questions.

A. Do responses made to self-referent items differ significantly from those made to matching other-referent items in terms of the amount of unfavourable attitudes, disturbances, and problems revealed?

B. Which type of item is more revealing of mal-adjustment?

C. Are both types of items "projective" techniques?

¹ Arthur C. Keating, M.A. Thesis, presented to the School of Psychology and Education of the University, Ottawa, Ontario, 1960, ix-130 p.

2.- The Study

The experiment was based on Allport's postulate that direct and projective measures of assessment would reveal the same information about the motives and attitudes of the normal person, but with the maladjusted individual, projective techniques would elicit repressed material which would be contrary to that expressed on direct measures of personality.

An experimental form of the Sacks' Sentence Completion Test, made up of sixty self-referent items and sixty matched other-referent items, mixed at random, was administered to 647 male subjects. Using the criterion of a psychiatric assessment, the experimental population was divided into an "adjusted" group and a "maladjusted" group. Fifty subjects were selected at random from the "adjusted" group to be compared to the twenty-nine in the "maladjusted" group.

Each set of sixty items constituted a "test", and each test was scored on a separate rating sheet. Each response was scored either "positive" or "negative" with reference to adjustment-maladjustment. "Negative" responses were summed to provide scores on the tests.

It was hypothesized that:

1. There would be no significant differences between the two forms of the Sentence Completion Test for either groups.
2. There would be no significant differences between groups on either form of the Sentence Completion Test.

3.- Results and Conclusions

Each group obtained a significantly higher mean score on the Other-Referent Form of the test. The two groups did not differ significantly on either Form of the test.

These results suggested that the other-referent type of item, by its nature, elicits a greater number of unfavourable responses. Neither one nor the other type of item would distinguish between the adjusted and maladjusted individuals. By the rationale on which the experiment was based neither type of item could be considered to be "projective".